



별첨1-4 AI 활용 보고서 (Track1)

1. 참가팀 정보

연구 내용	지정/선택 여부	지정	연구분야	바이오
	연구주제	ST + WSI 멀티모달 데이터를 이용한 tumor 예측 XAI 모델		

2. 활용 AI 모델 정보

AI 모델명	GPT 5.2	AI 모델 URL	https://chatgpt.com/
	Perplexity Sonar		https://www.perplexity.ai/
	Claude Sonnet 4.5		https://claude.ai/

※ 자체제작 AI 모델 활용 시, 재현성 검증을 위해 해당 AI 모델을 사용 또는 다운받을 수 있는 URL을 제시

3. 연구 절차별 AI 기여도(자체평가)

연번	연구 절차	배점 (A)	AI 기여도 자체평가 (0~100%) (B)	AI 기여도 산출 (C=A × B)	AI 작업내용 (간략히)
1	주제 선정 및 연구문제 도출	10	75%	7.5	후보 주제 목록 생성
2	선행연구 조사 및 문헌 검토	5	75%	3.75	관련 논문 요약 및 연구 흐름 정리
3	연구목적 및 문제 정의	15	50%	7.5	Pain point 및 연구 목표 정리
4	연구계획 및 방법론 설정	20	50%	10	모델 및 모듈 타당성 검토
5	자료(데이터) 수집	10	50%	5	후보 데이터셋 조사 및 데이터셋 개요 정리
6	자료(데이터) 분석	20	75%	15	코드 검토 및 실험 결과 분석 보조
7	논문 작성	5	50%	2.5	논문 초안 작성 및 번역
8	자체 리뷰 및 수정	15	75%	11.25	비문·오탈자 교정 및 연구 타당성 평가
총점		100		62.5	

※ AI 기여도 자체평가 기준: “의사결정 및 행동의 주체가 누구인가”

※ 항목별 배점 x 기여도 자체평가의 총합이 60% 이상이어야 함



불임

AI 작업내용 상세 (활용 증빙 스크린샷, URL, 로그파일 등)

1. 주제 선정 및 연구문제 도출

1) 연구 주제 후보 생성 및 구체화

- ChatGPT: <https://chatgpt.com/share/6954851d-608c-8007-8cd4-e29e47f99a96>

The image shows two side-by-side screenshots of the ChatGPT 5.2 interface. Both screens have a light gray background with a white input field at the bottom.

Left Screen (Thought for 32s):

- Header: ChatGPT 5.2
- Text area:
 - 할 것
 - 특히 ST 데이터 기반 tumor 예측 모델이 있으면 ST 데이터 기반이 아닌 모델보다 더욱 자세히 설명할 것
 - 모든 연구들을 참고 논문 링크를 반드시 첨부할 것 해 줘, tumor 예측이 아닌 다른 주제에 대해서도 위와 동일하게 해 줘.
- Message input: 좋아요, 면밀 전 '주제 fix'를 듣기 위해 (A) 주제 후보 지도를 먼저 펼치고, 이어서 (B) 'tumor 예측(종양 관찰)' 트랙과 (C) 종양 외(대안) 트랙으로 나눠 선별연구를 정리해 드릴게요. 각 연구는 '무엇/키워드/핵심원리/한계/개선점/링크' 순서로 아주 압축해서 적었습니다.
- Message input: A) 주제 후보 지도 (CV × Spatial Transcriptomics 관점)
 - T1. H&E ↔ ST 연계 "visual-omics" 파운데이션 모델: H&E 페치와 ST spot/유전자 발현을 동일 표현공간으로 놓고 대규모 사전학습(contrastive/CLIP)으로 표현학습 → 전이, 종양 예측/서브타입 예후 등으로 파인튜닝. [Nature](#)
 - T2. H&E만으로 ST의 유전자(경로) 예측: ST 비율 절감을 위해 histology→gene 매핑 (회귀/유니티스트스). 종양 마커, TME(면역세포 밀도) 프록시도 가능. [arXiv](#)
 - T3. GNN-Transformer로 ST의 공간구조 활용: spot/세포를 노드로 보고 공간그래프 +표현학습으로 분류/서브타입/진단. 종 ↓ 신경증양 등 실제 코호트 적용사례 존재. [Nature](#)
- Message input: + Message ChatGPT

Thought for 16s:

ST+이미지 multimodal tumor prediction하는 모델 연구를 진행한다고 할 때 pain point, 대상 contribution, 타당성 검토해 줘.

1) Pain point

- "정밀 과정"의 정의가 생각보다 어렵습니다.
- tumor prediction이라고 해도 무엇을 폐족하는지에 따라 티를 줄일 수가 있습니다.
- tumor vs normal은 초기 내 치료법 때문에 somewhat 더 잘 수행되는 경우가 있고,
- subtypes/graders에서는 비종 친자(슬라티드) 레벨 과정이어서 ST spot/유전자와 연결이 약한 역할(weak supervision)을 제가 합니다.

결국 '모델화하는 풍부한데(이미지+ST), 차별은 거칠다'는 구조가 생겨서 단순

- 1) ST 데이터를 활용한 multimodal model 연구를 목표하는 상황에서 주제 scope 확립을 위한 세부 주제 후보 선정을 위해 AI를 활용함.
- ST 데이터를 중심으로 computer vision과 biomedical AI를 결합할 수 있는 연구 방향을 제시하도록 지시하고, tumor detection/prediction, tumor subtype 분류, 예후 예측 등 가능한 세부 주제 후보들을 비교 및 정리함으로써 연구 scope를 단계적으로 구체화함.
- 나아가 ST 데이터를 활용한 computer vision 및 medical AI 관련 선행 연구 조사와 지시(2. 선행연구 조사 및 문헌 검토 단계와도 연결되는 내용)하여, 이를 바탕으로 본 연구가 기존 연구 대비 어떤 차별성과 확장 가능성을 가질 수 있는지에 대한 타당성 평가를 수행함.
- 배경 상황 설명과 함께 방향성의 큰 틀을 제시하고 그에 대한 세부 사항 및 근거 제공을 요청하였으므로, AI의 답변이 사용자의 의도 및 요구에서 크게 달라지지 않음.



2. 선행연구 조사 및 문헌 검토

1) 논문 review – Celcomen

- ChatGPT: <https://chatgpt.com/share/697a062d-f1fc-8007-9138-4623a89871fd>

The image shows two side-by-side screenshots of the ChatGPT interface. Both windows have the title 'ChatGPT 5.2' at the top.

Left Window: The title is 'Celcomen 논문 리뷰 쪽지'. It contains a PDF preview titled '2409-0150.pdf' with the following text:
Celcomen 논문 리뷰 쪽지. Celcomen이 해결하고자 하는 문제가 뭔지, 어떻게 해결하는지(복잡한 풍차를 생각하고 핵심 데이터에 중심으로), limitation이나 further work, 및 다른 점이.

아래는 Celcomen 논문에 대한 핵심 내용입니다. 주식 증명은 살피이고, 무슨 문제를 중요하게 보고, 어떤 아이디어로 문제를 해결하는지를 찾았습니다. (문제는 논문 본문에 근거합니다). □ celcomen

Right Window: The title is 'Celcomen의 핵심 아이디어를 우리 연구에 적용할 수 있는 방법' (Celcomen의 핵심 아이디어를 우리 연구에 적용할 수 있는 방법에 대해). It contains the following text:
Celcomen의 핵심 아이디어를 우리 연구에 적용할 수 있는 방법에 대해. 내 생각에는 Celcomen 보호 구조 자체를 사용하는 내용은 잘 알고, 필요하다면 explainability 강화에 해당 아이디어를 적용하거나 그래프 기반 fusion 등을 통해 진행할 수도 있을 듯.

좋은 성과 감사해요.
“Celcomen을 그대로 쓰지 않고, 아이디어를 풀어썼다”라는 판단이 지금 연구 스타일 (ST+WSI multimodal, ML explainability)에 잘 맞습니다.
마지막으로 Celcomen의 핵심 아이디어를 분석해 → 우리 연구에 적용 가능한 형태로 제작해서 배운데 불과요.

- ST 데이터에서 세포 내(intra-cellular) 및 세포 간(inter-cellular) 유전자 조절을 인과적으로 분리하고, 공간적 perturbation을 생성하는 선행 연구(Celcomen)에 대한 논문 리뷰를 지시하여, 논문의 문제 설정 및 핵심 아이디어를 파악하고 이를 본 연구에 적용할 수 있는 방안을 탐색함.
- Celcomen의 핵심 개념인 인과적 영향력(force) 관점을 attention 기반 explainability 프레임으로 재해석하여, 본 연구의 explainability 강화 및 fusion 전략에 적용할 수 있는 가능성을 검토함.
- 단순히 “논문을 review할 것”을 지시하기보다는 논문에서 중점적으로 볼 요소를 지정해 요청함으로써, AI가 사용자 친화적인 요약을 제공함.

2) 논문 review – scMODAL

- ChatGPT: <https://chatgpt.com/share/6969ef4a-7f34-8007-8d2d-825067b3a4f3>

1) Spatial Transcriptomics (공간전사체)

The screenshot shows two side-by-side ChatGPT sessions. The left session is about scMODAL, and the right is about scGNN. Both sessions include a PDF file and a summary of the model's strengths and limitations.

Left Session (scMODAL):

- PDF: 21587_029_80073.pdf
- Summary: 좋은 질문입니다. 이전 용어 정리 + 연구 포지셔닝을 정확히 해야 하는 자점이라서, 단정적으로 말씀드릴게요.
- Q: “scMODAL이 ST 기반 모델 중 current SOTA다” → 정확한가?
- A: ❌ 결론부터 말하면 그 표현은 암울히 틀립니다.
- scMODAL은 ST(Spatial Transcriptomics) 기반 모델의 SOTA가 아닙니다.
- 이유는 아주 명확합니다.
- scMODAL의 정체성
 - primary task: single-cell multi-omics integration
 - core setting: unpaired cross-modality alignment
 - ST는
 - 입력 modality 중 하나로 사용될 뿐
 - 주된 benchmark/task가 아님
- 논문에서도 scMODAL은 이런게 전의하니 ↓
- + Ask anything

Right Session (scGNN):

- PDF: 2310.02275v1.pdf
- Summary: 좋은 질문입니다. 우리 연구에서 MuSe-GNN은 핵심 아이디어를 적용할 수 있는 장점은?
- Q: 좋은 질문입니다.
- MuSe-GNN은 “모듈 문제를 그대로 브이자면 아이디어 단위로는 제 몫 드린 게.”가 명확한 논문입니다.
- 우리 연구방법 + WSI multimodal tumor prediction / XM (제형) 예측에서 적용 가능한 핵심 아이디어를 구조화로 정리해 놓았습니다.
- 아래는 “그래도 가져보면 한 번은 것 / 가져오면 감동할 것”을 분류해서 설명하겠습니다.

- Multi-omics 데이터 통합에서 강한 feature correspondence 없이도 alignment를 수행하는 선행 연구(scMODAL)에 대한 논문 리뷰를 지시하여, 논문의 문제 설정 및 핵심 아이디어를 파악하고 이를 본 연구에 적용할 수 있는 방안을 탐색함.
- ST를 종양 예측의 직접적인 입력보다는 WSI와 scRNA를 연결하는 중간 표현(bridge modality)으로 활용하는 관점을 도출하였으며, ST의 좌표 정보를 구조적 제약으로 활용하여 공간적으로 인접한 spot들의 embedding은 서로 가깝게, 멀리 떨어진 spot들의 embedding은 멀어지도록 정규화하는 방식이 효과적일 수 있음을 논의함.
- 논문의 내용뿐 아니라 우리 연구와 연결할 수 있는 요소를 질의하여, AI를 “챗봇”을 넘어선 “연구동료”로 활용함.

3) 논문 review – MuSe-GNN

- ChatGPT: <https://chatgpt.com/share/697abe80-0288-8007-a34a-88e7f4c5d992>

The screenshot shows two side-by-side ChatGPT sessions reviewing the MuSe-GNN paper. Both sessions include a PDF file and a comparison table of the research question and MuSe-GNN's approach.

Left Session (MuSe-GNN):

- PDF: 2310.02275v1.pdf
- Summary: MuSe-GNN 논문 리뷰에 질. MuSe-GNN이 해결하고자 하는 문제가 뭘지, 어떻게 해결했는지(복잡한 공식은 생략하고 핵심 아이디어 중심으로), limitation이나 further work가 있다면 뭡지.
- Q: MuSe-GNN 논문은 핵심 아이디어 중심으로 정리해서 리뷰드리겠습니다. (복잡한 수식은 생략하고, 왜 이 문제가 생기고 → 왜 이런 설계를 했는지' 위주로 설명하겠습니다)
- 1. MuSe-GNN이 해결하고자 하는 문제
 - 문제 배경
 - 최근 생물학에서는
 - scRNA-seq, scATAC-seq, spatial transcriptomics 등
 - 모달리티-조직-종(species)이 서로 다른 대규모 멀티모달 데이터가 빠르게 늘고 있습니다.
 - 기존의 멀티모달 통합 방식은 주로 “cell-centric integration (cell을 anchor로 정렬)”에 의존했는데, 대. 논문에서는 이 접근이 근본적으로 한계가 있다고 지적합니다.
 - spatial data에서는 한 spot이 여러 cell ↓ 혼합 신호임
 - modality마다 관측 단위와 noise 특성이 다름

Right Session (Comparison):

① 전제: 우리 연구와 MuSe-GNN의 차이		
항목	MuSe-GNN	우리 연구
기본 단위	gene	patch / patch / gene
그래프 노드	gene only	root, patch, gene (질병적으로 heterogeneou)
복잡	gene embedding (dataset, agrégat)	tumor prediction + explainability
superior	fully unsupervised ↓	weakly supervised (slide-level)

- 유전자를 공통 언어로 삼아 서로 다른 multimodal biological graph를 정렬하는 선행 연구 (MuSe-GNN)에 대한 논문 리뷰를 지시하여, 논문의 문제 설정 및 핵심 아이디어를 파악하고 이를 본 연구에 적용할 수 있는 방안을 탐색함.



- 이를 바탕으로, 본 연구의 ST-WSI multimodal 설정에서도 단순한 feature 결합을 넘어 생물학적 의미 단위의 정렬을 loss 수준에서 유도하는 설계 가능성을 탐색함.
- 또한 scMODAL 논문과 동일한 맥락에서, 공간 좌표 정보를 입력 feature로 직접 사용하기보다는 spot 간 관계를 정의하는 relation generator로 해석하는 관점을 도출하였으며, 이를 통해 좌표 정보를 기반으로 ST 표현 학습 및 explainability를 강화하는 방향을 연구 설계에 반영하고자 함.



3. 연구목적 및 문제 정의

1) 연구 목적, 문제 정의 및 연구 novelty 파악

- ChatGPT: <https://chatgpt.com/share/697ac127-e958-8007-8ee1-9e66a0378283>
- Perplexity:https://www.perplexity.ai/search/daehoe-jeculyong-nomnuneul-jag-u_8rk6MR.OUhAe8bbW78Q#4

The image shows two side-by-side screenshots of ChatGPT instant responses. Both screenshots have a header "ChatGPT 5.2 instant" and a "Share" button.

Left Screenshot (WSI+ST Multimodal Research):

- Text: "자본적으 나는 대화를 바탕으로, 본 연구의 목적과 문제 및 학습률 향상의 관점에서 차이가 있는지, 특히 연구에서 이를 어떻게 배웠는지 하는지"
- Text: "질문입니다. 지금까지 논의한 모든 예제, 데이터 개요, 실험 방법을 종합해서 “본 연구의 목적과 문제 및 학습률 향상”을 논문에 바로 들어갈 수 있는 수준으로 글라써 보겠습니다. (서술 형식으로, 필요하면 이를 문단 분리도 가능합니다.)"
- Text: "1. 문제 정의 (Problem Definition)
최근 생체 영상(WSI)과 같은 전시체 데이터(Spatial Transcriptomics, ST)를 결합한 단독 모델 학습률 향상을 위한 특수한 특성을 동시에 조작할 수 있다는 점에서 주목 받고 있습니다. 그러나 현재 해당 분야에는 다음과 같은 구조적인 원인이 존재합니다.
첫째, WSI-ST 학습모달 모델의 대부분을 담는 월드로드 설계되어 보면, 모델의 학습률 향상 과정에 대한 해석 가능성(explainability)이 충분히 확보되지 못한다는 점입니다. 물론 기존 연구는 이미지 라벨 단위 또는 t-SNE 단위의 feature를 단순히 aggregate하여 slide-level prediction을 수행하는 데 그치며, 어떤 공간적 영역이나 어떤 유전자 신호가 예측에 기여했는지에 대해서는 명확한 설명을 제공하지 못합니다. 이는 의료 도메인에서 요구되는 신뢰성, 재현성, 일상적 해석 가능성 축면에서 중요한 제약으로 자족합니다.
둘째, ST 데이터를 다루는 기존 모델들은 고수준의 특성을 활용한 보조 정보로 사용하거나 예제 배운 후, scRNA-seq와 유사한 형태로 처리하는 경우가 많다는 문제점이 있습니다."

Right Screenshot (scRNA-seq Research):

- Text: "scRNA 기반으로도 동사형 연구가 있을 것 같았는데, WSI+ST은 어떠한가요? WSI와 multilevel로 실험하는 것의 차이나 novelty가 있을까요?"
- Text: "좋은 질문입니다. 이런 심사차 질문에서 거의 확실히 나올 질문예고, 지금 연구의 정밀성을 청취해 주는 게 중요합니다.
질문보다는 말씀입니다. 단, “그가 했다” 자체적으로는 부족하고, 어떤 문제를 해결해 마무리하기 때문입니다. 아래를 논문용 논리로 정리해 보겠습니다."
- Text: "6. 전체 정리: scRNA 기반 연구가 어떤 차이를 가져온가?
글습니다.
• scRNA 기반은 암 예측, subtype 분류, prognosis 예측 연구는 90% 학부 참고
• WSI + scRNA는 sample-level에서 결론하는 연구도 일부 존재합니다.
하지만 이를 연구는 결론적으로 다음 가정을 들립니다.
scRNA는 “공간 정보는 계기화 분자학 표적”이며,
WSI의 관계는 slide-level 또는 patient-level에서는 정의된다.
이 가정이 바로 ST를 쓰는 순간 깨집니다."

- 주제 선정과 선행 연구 분석 대화를 기반으로, WSI+ST multimodal 연구의 목적과 novelty를 구체화함.
- ST 데이터를 활용함으로써 유전자가 “어디에서 발현되었는지”에 대한 공간 정보를 유지할 수 있으며, 이는 공간 정보가 소실되는 scRNA-seq 기반 기존 연구 대비 중요한 강점임을 확인함.
- ST 단독 모델에서는 조직학적 morphology 정보를 직접적으로 활용하기 어렵다는 한계를 WSI와의 multimodal 구조를 통해 보완할 수 있음을 정리하였으며, WSI는 조직 구조와 형태학적 패턴을, ST는 분자적 신호를 제공함으로써 두 modality의 결합이 강점을 가짐을 확인함.
- 본 연구의 목적을 단순한 암 존재 여부 예측을 넘어, 모델의 예측이 어떤 공간적 영역과 어떤 분자 신호의 결합에 의해 이루어졌는지를 구조적으로 추적하고 설명하기 위한 합리적인 연구 파이프라인 및 모델 구조를 고안하는 것으로 설정함.



4. 연구계획 및 방법론 설정

1) 모델 구조 검토

- ChatGPT: <https://chatgpt.com/share/697afdd8-f938-8007-9196-3f57c64598be>

Left Model (Multi-task Learning):

ST-WSI 연구에서 위와 같은 구조로 모델 아키텍처 구성을 했는데, 각각의 모델은 해당 데이터에 맞는 구조를 갖습니다. 예전에는 sanity check 등 모델 두 가지가 XAI 모델과 주제 모델로 구분되었지만, 최근에는 두 가지 모델을 통합하여 한 번에 이미지를 XAI 모델로 활용하는 방식으로 전환되었습니다.

이전에는 두 가지 모델을 각각의 모델로 구현하였지만, 최근에는 두 가지 모델을 통합하여 한 번에 이미지를 XAI 모델로 활용하는 방식으로 전환되었습니다.

Right Model (Downstream Task-oriented):

downstream task마다 다른 구조를 어떻게 수립하게 될까요?

다른 질문입니다. 예전에는 두 가지 모델을 각각의 모델로 구현하였지만, 최근에는 두 가지 모델을 통합하여 한 번에 이미지를 XAI 모델로 활용하는 방식으로 전환되었습니다.

이전에는 두 가지 모델을 각각의 모델로 구현하였지만, 최근에는 두 가지 모델을 통합하여 한 번에 이미지를 XAI 모델로 활용하는 방식으로 전환되었습니다.

다른 질문입니다. 예전에는 두 가지 모델을 각각의 모델로 구현하였지만, 최근에는 두 가지 모델을 통합하여 한 번에 이미지를 XAI 모델로 활용하는 방식으로 전환되었습니다.

- 본격적인 모델 구현 전 구상한 모델 architecture가 타당한지 검토하고, 구현 세부사항 및 확장 가능성을 검토함.

2) CLIP 방법론 차용

- ChatGPT: <https://chatgpt.com/share/696d9bd7-e1a8-8007-91c7-451ebb6c9c5f>

Left Diagram (Global View):

CLIP은 이미지 + 텍스트 + 멀티모달 데이터를 각각 학습한 후, 같은 음성Embed 공간으로 전개 후 영상과 대비학습으로 학습하는 구조입니다. 아래에서 구조 → 차 차별화된 학습을 동시에 실현합니다.

1. 전체 구조 한눈에 보기

CLIP은 두 개의 인코더를 별개로 합니다.

- Image Encoder: 이미지 → 이미지 표현 헤드
- Text Encoder: 텍스트 → 텍스트 표현 헤드

각각의 출력은 같은 차원의 멀티모달 벡터로 조건으로 신경�음(contrastive loss)으로 학습됩니다.

Right Diagram (High Level Architecture of CLIP):

우리 연구에서는 computational resource 제한 때문에 인코더를 학습시키는 게 어려워서 인코더 뒤에 linear proj레이어를 추가하고, 인코더는 freeze한 뒤 proj레이어만 학습하는 방식으로 진행하고자 하는데 CLIP에서 사용하는 접근이 이게 맞지? 인코더는 freeze하고 레이어만 학습하는 거

질문부분에 분명히 말씀드렸다면 “CLIP에서 실제로 한 것”과 “지금 연구에서 하려는 설정”은 정확히 같지는 않지만, 개념적으로는 CLIP-style로 정의화 가능한 접근입니다. 다만 논문에서의 CLIP과 동일하다고 말하면 안 되고, CLIP의 projection-head 분리 철학을 사용했다고 표현하는 게 정확합니다.

아래에서 CLIP이 실제로 한 것 → 지금 설정과의 관계 → 논문에서 어떻게 말해야 하는지 순서로 설명드리겠습니다.

1. CLIP에서 실제로 한 것 (팩트 정리)

CLIP은 인코더를 freeze하지 않았습니다

CLIP에서 실제로 한 것 → 지금 설정과의 관계 → 논문에서 어떻게 말해야 하는지 순서로 설명드렸습니다.

- 이미지-텍스트 multimodal 모델인 CLIP의 encoder-proj 구조를 본 연구에 적용하기 위해
 - (a) CLIP 원 논문을 기반으로 모델 구조 및 학습 방식의 설계 의도를 검토하고, (b) CLIP 모듈 중 linear projection layer의 역할과 필요성을 중심으로 우리 연구에 차용 가능한 요소를 파악하였으며, (c) 본 연구의 계산 자원 제약을 고려하여 encoder를 freeze한 상태에서 proj layer만 학습하는 접근이 타당한지에 대한 분석을 수행하도록 함.
- 이를 통해 CLIP의 end-to-end 학습 방식과 본 연구의 제한 조건 하에서의 변형된 적용 방



식 간의 차이를 명확히 구분하고, 본 연구 설정이 CLIP-inspired downstream adaptation으로 정당화될 수 있음을 확인함.

3) 모델 코드 수정 검토

- Perplexity:<https://www.perplexity.ai/search/meoltimodalin-ai-modelyi-kodeu-hnFTaBFwSRG6FPCoyyJUUQ>

The image shows two side-by-side screenshots of the Perplexity AI interface. Both screenshots display code snippets and their analyses. The left screenshot is for a 'BERT' model, and the right one is for an 'ST' model. Each screenshot includes a header with 'Answer', 'Links', 'Share', and 'Download Content' buttons. The main content area contains the code snippet, its analysis, and a 'Show more' link. The right screenshot also includes a 'Reviewed by 16 source(s)' section and a note about 'concat' being used in the ST module.

- Multimodal로 구현된 모델을 unimodal로 사용할 수 있도록 코드를 수정하기 위해 구현 방식 개요와 예시 코드를 요청함.

4) Encoder architecture 검토

- ChatGPT: <https://chatgpt.com/share/697b0be1-ff60-8007-b56f-28ea1cd3b0f5>

The image shows two side-by-side screenshots of a ChatGPT conversation. The left screenshot is titled 'ChatGPT 5.2 instant' and discusses the differences between 'scBERT' and 'scGPT', specifically mentioning 'token-level dependencies' and 'cross-task learning'. The right screenshot is also titled 'ChatGPT 5.2 instant' and discusses the performance of 'scBERT' and 'scGPT' in 'single-cell RNA' tasks, noting that they 'don't work well' in these tasks.

- ST 인코더에 scBERT와 scGPT 중 어떤 모델을 사용하는 것이 나은지 검토하기 위해, 각 모델의 장단점과 본 연구와의 적합성 분석을 요청하였고, BERT와 GPT 구조적 특성과 사전 학습된 parameter 접근 여부, 구현 난이도 등을 고려해 각 모델의 적용 가능성을 검토함.



3) 모델 코드 검토

- ChatGPT: <https://chatgpt.com/share/697afb32-4614-8007-adc7-f93159051df4>

The screenshot shows two code snippets side-by-side. On the left is `model.py` and on the right is `ablation.py`. Both snippets include Korean annotations explaining their purpose and how they interact with the system.

```

model.py 코드 간략화 퀘
이제까지 모델 구조도가 잘 구현되었는지 확인하고 fusion
을 일부선 품분 예전에 둔 거니까 concat 알고 다른 걸로
있는 게 정상. 학습률을 맞추거나 gradient 폭주가 일어날
만한 치명적인 오류가 있는지 계크해. minor한 최적화 문
제 경고는 무시해.

줄은 코드입니다. 현재 구조는 볼루린 모델 디터그램과 개별학습을 잘 일치하고, 학습
이 애에 맞춘다 gradient가 더진다 수준의 차별적인 설계 오류는 1개단 영향하게 있고,
나머지는 의도에 의해 일부 예전에 품분 혹은 학습률을 정했습니다.
아래에서 (A) 구조 정합성 → (B) 치명적 버그 → (C) gradient/학습 정합성 관찰 리스너
순서로 보여드립니다.

```

A. 모델 구조도와의 정합성 체크

1. 전체 파일 프리미
...
+ click anything

```

ablation.py 코드 간략화 퀘
모델 구조도와 맞는지 확인합니다. 현재 구조는融混된 모델 디터그램과 개별학습을 잘 일치하고, 학습
이 애에 맞춘다 gradient가 더진다 수준의 차별적인 설계 오류는 1개단 영향하게 있고,
나머지는 의도에 의해 일부 예전에 품분 혹은 학습률을 정했습니다.
아래에서 (A) 구조 정합성 → (B) 치명적 버그 → (C) gradient/학습 정합성 관찰 리스너
순서로 보여드립니다.

```

B. 구조 정합성

설정	img_encoder	st_encoder	value	spot_embeds	총 학
image+id	o	o	o	labelizing	정상
image only	o	x	x	img_text	정상
st only	x	o	x	st_text	정상

C. 치명적 오류

아래는 `major 오류만` 등록합니다.

- 모델 architecture 구현 이후, multi/unimodal ablation을 위한 encoder on/off 구조가 의도대로 동작하는지 코드 레벨에서 검토를 요청하여 실험 분기(WSI only/ST only/WSI+ST)가 논리적으로 올바르게 구성되었음을 확인함.
- 치명적인 런타임 오류를 사전에 식별하고, 최소 수정으로 모든 ablation 설정에서 안정적으로 학습 및 추론이 가능하도록 개선함.
- 모델 학습에 긴 시간과 연산 자원이 소요되는 상황에서 학습 도중에 문제가 발생하면 치명적이고, 긴 코드를 수동으로 디버깅하는 것은 비효율적이므로 AI를 통한 코드 검토가 필수적임.

3) train/test 코드 검토

- ChatGPT: <https://chatgpt.com/share/697af85f-ec9c-8007-87b9-47f49065cec9>

The screenshot shows two code snippets side-by-side. On the left is `train.py` and on the right is `test.py`. Both snippets include Korean annotations explaining their purpose and how they interact with the system.

```

train.py 코드 간략화 퀘
img-only / st-only / img+st
수집한 코드인데, 유통 않는지 전제적으로 걸트해. 학습이
간 불 칸은 오류(학습 불가). gradient 폭주 등 major한
오류가 있으면 알려주고, 그냥 식소한 최적화 문제같은 건
우리에게.

```

전체적으로 구조는 잘 짜여 있고, 의도와는

- * st-only / img-only / img+st
- * 가지 ablation이 train.py에서 딸길하게 분기되도록 잘 구현되어 있습니다.
- * 학습하는데 있어서 gradient 폭주 같은 major 오류(학습 불가)에 맞춰서 브레이크 탈출하는 부분은 없습니다.
- 다면 “설정에 학습이 안 되거나 결과가 이상해질 수 있는 “중요한 포인트 3개”라는 반드시 첨고 남아야 합니다
- (그 외 식소한 최적화/스크립트 문제는 코드대로 무시하겠습니다.)

결론 요약

- * ❌ 즉시 학습이 제자리는 빠그. 오류
- * 🔍 설정에 따라 학습이 “의미 없이 될 수 있는” 구조적 문제로 있을 (애매한)
- * 🔍 gradient 폭주를 직접 유발하는 코드 없음 (e.g. 잘 들어가 있음)

Major Issue 1. `accue_steps + batch_size=1` 조합에서 마지막

test.py (ablation 포함)

- * st-only / img-only / multimodal 모두 정상적으로 학습 가능
- * gradient 폭주 저단위는 곳 몇몇
- * AMP / grad clipping / optimizer step: 로직 논리적으로 안전
- * acc가 안 고를 수 있는 설정 레인지(예: 배그는 예상)

“코드 때문에 학습이 안 된다”는 보통은 없음

test.py (XAI + sample-level 예측)

- * forward / unpack / 처리 로직 정립
- * mlm_mlm / gen_mlm 같은 대로 연관해 skip
- * mlm_mlm / gen_mlm 같은 대로 연관해 skip
- * mlm_mlm / gen_mlm 같은 대로 연관해 skip

- 모델 학습 및 평가 코드(`train.py`, `test.py`, `train_ablation.py`, `test_ablation.py`)를 점검하기 위



해 ablation 설정(st-only/img-only/multimodal)에 따른 로직 분기, gradient 흐름, 학습·추론 안정성 등에서 치명적인 오류가 없는지 검토함.

- 코드 리뷰를 통해 test 단계는 XAI(UMAP 시각화, 중요 patch/gene/spot 추출)에 집중하고, 성능 지표(ROC/AUC, P/R/F1)는 train/validation 단계에서만 계산하는 역할 분리를 명확히 정리함.

3) 성능 개선 전략

- ChatGPT: <https://chatgpt.com/share/697af50b-5618-8007-a090-2055454953c4>

The image shows two side-by-side screenshots of the ChatGPT 3.2 interface. Both screenshots have a light gray background with a white header bar containing the ChatGPT logo and a share button. The left screenshot contains text from a user about a 12-layer fusion model and its cross-entropy loss function. The right screenshot is a response from ChatGPT. Both screenshots include a bottom navigation bar with icons for back, forward, and search.

Left Screenshot Content:

• 모델: 교재(DNA)를 넣고 fusionation 모델
• Loss function 수정(현재 cross-entropy 손실로 사용하고 있음 대신 img-ST으로 커스텀하는 것)
• 풀이가 있어

좋습니다. 최근 커널 12개(구조-fusion) 풀을 SDM 성능에서 소프트에서 "가장 혼란한 베이스라인"을 먼저 고르는 한편이면 그 다음 성능 개선은 대체로 마련 속으로 나옵니다.
(정상 적용 신이죠/효과가 큰 것부터 위쪽에 두겠습니다.)

1) 데이터/타일링/샘플링 전략 (WSI 속 성능이 가장 많이 향상되는 구간)
• 파일 풀기 필터링, 배경/환경/글자/디자인 제거, 조작 퍼스팅 기반 파일 주제.
• Top-K 비슷한 방식 고도화
• 현재 템포/글로벌 상용화다면 → 예측에 기여하는 타일원의 확장 브로드 하는 방식 (예: attention score 기반 hard mining).
• 패러스 캐싱 타일원 같은 위치를 서로 다른 래플(예: 5x/10x/20x)로 놓고 concat/intersection으로 통합.
• WSI augmentation 차작성: 색상(3d, stain) variation이 고려 stain jitter / Macenko/Reinhard normalization 같은 색상 정규화가 물론 baseline을 크게 끌어 올립니다.
• 플래스 풀교환 대용: organ-level/file-level 관계 중요한 weighted sampling 또는 class-balanced sampler.

+ do anything

Right Screenshot Content:

• mRNA와 ST의 가장 큰 차이는 좌표 정보의 투부라고 생각해서, 좌표 기준으로 가까운 spot의 임베딩은 가까워지게 → 먼 spot의 임베딩은 멀어지게 구현하는 방식을 생각 중이야.

이 접근이 탐색원지 한 번 검토해 보고, 이를 구현한다면 어떤 도움이 어떤 식으로 구현해야 하는지 전체 코드는 주지 않고 대략적으로 알려 줍니다.

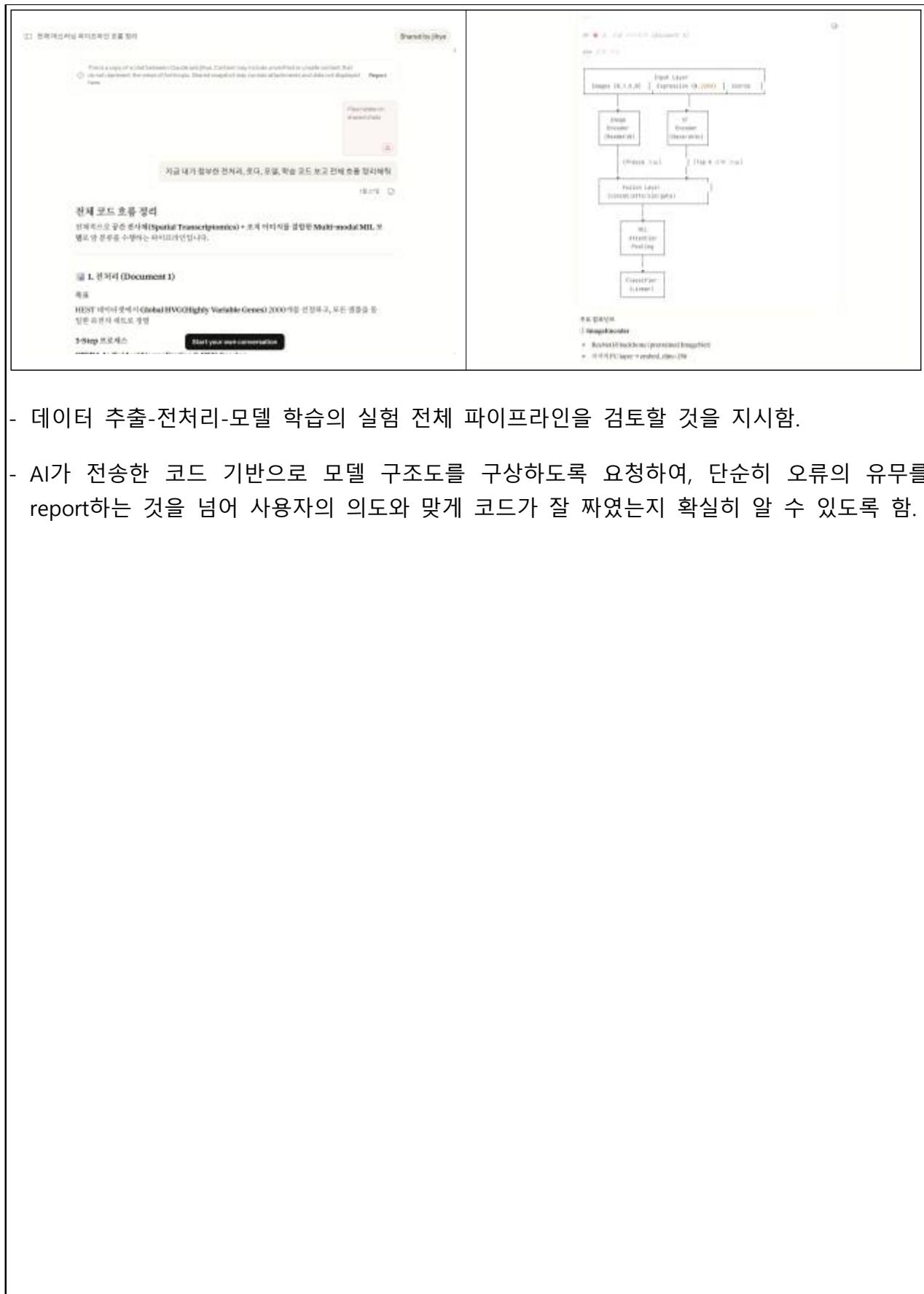
마땅합니다. 다만 “좌표만으로 가까우면 절대 무조건 가까워야 한다”를 너무 강하게 말면, 영/광상 경계(내부/외부 경계)처럼 계약율에도 다른 상태인 경우를 봄에서 성능이 떨어질 수 있어요. 그래서 결론은:
• 좌표 기반 균형성 정규화는 좋은 보조 신호가 될 수 있음
• 대신 (i) 토큰에서만, (ii) 사용자를 위해, “너무 거리워도 다른 표현을 해용”하는 형태(마진/양안한 카울) 등 두는 게 안전합니다.
마리는 “왜 필요하가 → 어떻게 설계하는가 → 어디에 넣는가 → 잘 스냅 시뮬레이션” 흐름으로 정리해 드릴게요.

1) 왜 이게 도움이 되나?
ST는 spot이 2D 그리드/좌표 위에 있고, “조직 구조는 조직적으로 연속성(smoothness)”을 갖는 경우가 많습니다. 즉,
• 다른 spot들은 비슷한 cell-type mixture, microenvironment을 공유할 확률이 높고,
• + do anything

- Baseline 모델 학습 이후, 지금까지 논의한 내용(논문, 기존 구현 사항 등)들을 고려하여 모델 성능을 개선할 수 있는 여러 전략들을 요청해 가능한 방안을 검토함.
- ST 데이터의 좌표 정보를 활용한 보조 loss 설계 방안을 검토하기 위해, AI를 활용하여 좌표 기반 임베딩 정규화(근접 spot 간 유사도 강화, 원거리 spot 간 분리) 접근의 타당성과 잠재적 한계를 분석함.
- 해당 보조 loss를 기존 cross-entropy 기반 이진 분류 학습에 결합하는 방식에 대해, 어떤 표현 수준(spot-level embedding)에 적용하는 것이 적절한지와 모듈 삽입 위치(ST 인코더 출력 이후)를 중심으로 구현 전략을 구체화함.

4) 전체 머신러닝 파이프라인 정리

- Claude: <https://claude.ai/share/0807d86d-4ec5-44a2-bf1a-f488807e0352>





5. 자료(데이터) 수집

1) 데이터 조사 및 선정

- ChatGPT: <https://chatgpt.com/share/697a0036-1fc4-8007-aa55-3de33b06d543>
- Perplexity: <https://www.perplexity.ai/search/st-spatial-transcriptomics-wsi-SDQTrPOSlyNZk0Y6fu0Hg#1>

The image shows two side-by-side screenshots of online search results.

Left Screenshot (ChatGPT):

- Header: ChatGPT 5.2
- Text: ST(Spatial Transcriptomics), WSI(Whole Slide Image) 를 모두 포함하고 ST의 각 스팟이 WSI에 매칭되어 있거나 매칭 가능한 정보가 들어있는 데이터셋을 알려 줘요
- Text: 3 steps completed >
- Text: ST와 WSI가 매칭된 데이터셋으로는 HEST-1k와 STImage-1K4M이 대표적입니다. 이 데이터셋들은 Spatial Transcriptomics(ST) 스팟의 gene expression과 해당 위치의 Whole Slide Image(WSI)를 페어링하여 제공하며, 다수의 연구에서 multimodal 모델 학습에 활용됩니다. [nature]
- Section: HEST-1k
- Text: 1,229개의 ST 프로파일과 해당 H&E WSI를 포함하며, 26개 장기, 인간/마우스 종, 26종 암 타입을 커버합니다. 각 ST 스팟은 WSI 내 위치와 정렬되어 있으며, Hugging Face에서 다운로드 가능 (전체 1TB 이상). Scampy 형식과 패치 이미지, 메타데이터가 제공되어 bioinformatics 파이프라인에 적합합니다. [huggingface]
- Section: STImage-1K4M
- Text: 1,149개의 ST 슬라이드(400만+ 스팟)로, 각 스팟의 고해상도 histopathology 이미지 타일과 15,000-30,000 차원 gene expression을 매칭합니다. 11개 조직 타입, 3종 ST 기술(Visium 포함)을 지원하며, GitHub에서 접근 가능합니다. Spot-level 정밀 정밀 포함되어 ST-WSI fusion 모델에 이상적입니다.
- Text: (1) Viewing a shared thread. Your follow-ups will be private to you.
- Text: Ask a follow-up
- Buttons: Search, Refresh, Home, Stop, Up, Down

Right Screenshot (Perplexity):

- Header: ChatGPT 5.2
- Text: HEST, STImage 데이터셋 각각 간단히 줄여서 줘 데이터 규모, tissue, disease state, organ 등너그 회복 의사할 필요면 알려줘
- Text: 알겠습니다. HEST와 STImage를 연구자가 바로 찾을 수 있게 핵심만 비교 요약해 드리겠습니다.
- Text: (우리는 논문·공개 분석·기준화 대학적 규모이며, 능형 관점에서 중요한 모인트 취주법입니다.)
- Section: HEST (Human spatial transcriptomics atlas)
- Text: * 헌 출 오류
- Text: WSI + ST(주로 Visium) 데이터를 대규모로 모아둔, 잘 중심화된 표준화된 benchmark
- Section: 데이터 규모
- Text: * 채택 수 ~100개 내외 (예전 몇몇 기관은 미국 디스 배종)
- Text: * 스팟 수: 성별 및 구천 개 (Visium standard)
- Text: * 유전자 수: 수천~수만 (예전 iHVG dataset 사용)
- Section: Scope
- Text: * Species: Human only
- Text: * Disease state:
- Text: + 조직 분류(위장)

- ST 데이터 중에서도 암 관련 연구에 활용 가능한 데이터셋을 중심으로, Hugging Face 등 플랫폼의 데이터셋이나 raw 데이터 수준에서 접근 가능한 공공 repository 조사를 요청함.
- 이후 본 연구에서 활용할 HEST-1k 및 STImage-1K4M 데이터셋에 대해, 데이터의 규모, 포함된 modality(WSI, spatial gene expression), species, organ 범위 등 연구 설계에 필요한 최소한의 메타 정보 요약을 지시함.



6. 자료(데이터) 분석

1) 논문 review – HEST-1k

- ChatGPT: <https://chatgpt.com/share/697aeeae-7770-8007-bd15-6ac18b4438ef>

The screenshot shows two side-by-side ChatGPT instances. The left instance is titled 'ChatGPT 5.2 [instant]' and displays a PDF titled '2406.16192v1.pdf' which is the HEST-1k dataset. It contains text about the dataset being a collection of gene expression, metadata, and history from various sources. The right instance is also titled 'ChatGPT 5.2 [instant]' and discusses the 'HEST-1k benchmark' which uses UNIL CONCH as its foundation model. It highlights that the benchmark uses a multi-modal approach combining ST and WSI data.

- 데이터셋 선정 이후, 주요 실험 데이터셋으로 HEST-1k를 채택함에 따라 해당 데이터셋의 논문에 대한 체계적인 리뷰를 지시하여, HEST-1k가 기존 ST 데이터셋과 비교해 가지는 규모, 구성, multimodal(WSI-ST) 특성 등 정보를 파악하고 본 연구에의 적합성을 검토함.
- 이때 논문에서 제시하는 Dataset, Library, Benchmark의 세 가지 구성 요소를 구분하여 분석하였고, 데이터셋으로서의 특징뿐 아니라 데이터 접근 방식(공식 라이브러리)과 실험 설정(benchmark)을 종합적으로 이해하고자 함.
- 또한 논문에 포함된 benchmark 실험을 중심으로 실험 설정과 baseline 모델 구성, 그리고 ST-only, WSI-only, ST+WSI 조건에서의 성능 스케일을 분석하도록 질의함.

2) 논문 review – STImage-1K4M

- ChatGPT: <https://chatgpt.com/share/697af27e-727c-8007-95af-459f966806ff>

The screenshot shows two side-by-side ChatGPT instances. The left instance is titled 'ChatGPT 5.2 [instant]' and discusses the STImage-1K4M dataset, mentioning it includes gene expression, spatial transcriptomics, and auxiliary data. The right instance is also titled 'ChatGPT 5.2 [instant]' and explains the 'sub-tile' concept, defining it as a spatial transcriptomics gene expression patch of size 1:1. It also provides context about WSI and image-gene pairs.



- 보조 데이터셋으로 STImage-1K4M을 검토함에 따라, 해당 데이터셋의 논문에 대한 구조적 리뷰를 지시하여 STImage-1K4M의 특성을 파악하고 본 연구에서의 활용 가능성을 검토함.
- 논문에 포함된 contrastive learning 기반 실험을 중심으로, image encoder fine-tuning 방식과 평가 설정을 검토해 본 연구에서 HEST-1k를 주요 실험 데이터셋으로 유지한 채 STImage-1K4M을 보조적 학습 데이터로 사용하는 전략의 타당성을 검토함.

3) 코드 검토를 통한 실험 환경 설정

- ChatGPT: <https://chatgpt.com/share/6969b0b1-af80-8007-b33a-fc4ec2d44f5d>

train.py 코드입니다. 모델 충전: safety_a, bfusion 품션: concat, att, sim, gate 조합으로 본업해서 모델 train을 할 때 일정 간 통일해야 하는 부분 정리해 주세요.

질문들이 (모델 충전: safety/a/b) + [fusion: concat/att/sim/gate] 조합으로 나눠서 train/test를 때, 결과를 긍정하게 비교하려면 '출석체크' 대로 나누자는 전전적 통일'에 합니다. 아래 항목들을 잘 글을 규약으로 고정해주는 게 핵심입니다.

1) 데이터/현처리: 옵션과 무관하게 반드시 고정할 것

(A) train/test split 기준

- 블록 단위 splitWS 단위(단지 훈련이 끝나면 내부 spot을 살피고는 고정)
- split 파일을 통해 기장해서 광동(예: splitA/train-test, splitB/test.txt)
- (주의) HEST처럼 원자/케이스 충돌이 있을 수 있으면 patient-level split 어두운 경지

(B) label 배정 규칙

- max_squared_l2_norm에 무언가(예: normal-D or multi-hot) 형 태형으로 고정
- crossEntropyLoss, 기준이므로 label dtype/shape도 동일해야 함 (예: float32 이나 float64)

(C) 입력 spot 세팅 규칙

- float32

그냥 config 파일 하나 따로 다시 정리하는 게 낫겠지만!

네, 그게 제일 좋습니다.
지금 상황에서는 거의 충수에 기겠습니다.
다만 '아무 config' 하나 대 번들다가 아니면 맥락이 명확히 분리된 config어야 될 분야에 한 해둡니다. 이해차원 청취하시는 걸 합니다.

왜 config 하나로 묶는 게 맞나
지금 코드 기준으로 보면,

- 쪼글 훈련(splitW / a / b)
- fision 충전(concat / att / sim / gate)
- data split / top-k gene / spot sampling
- freeze 예후 / AMP / accumulation

이게 코드 예제처럼에 합쳐져있으면 좋겠다는 것입니다.
이 상태에서 뜯을면 100% 이런 일이 생깁니다.

- "나는 eval()인데 왜 acc가 다른지?"
- "n은 안 바꿨는데 결과가 왜 다르지?"
- "gamma가 대로하게 다른데..."

⇒ config 단일 파일 single source of truth 없으면 비교 실질 의미가 사라집니다.

- Option별 모델 성능 테스트 시 팀원들 간 실험 환경 통일을 위해 변수 설정을 요청해 어떤 변수를 통일해야 하는지, 다르게 두어도 되는 설정은 무엇인지 파악함.

4) 학습 로그 분석 및 실험 코드 오류 해결

- ChatGPT: <https://chatgpt.com/share/697b049d-150c-8007-a56a-c39153f1375f>

Training: 896/1925 batches/epoch, 2.375s/it, acc=89.9%, loss=0.34720

평화도가 저마다 높게 나오는데 획기 문제일까? 학습 코드에는 문제 있는 걸 확인했다.

로그만 보면 '모델이 너무 잘 맞았다'고 보다는, 데이터/평가 쪽에서 강하게 의심해야 할 진짜가 아니라 게 겹쳐 있습니다.
학습 코드 자체에 버그가 없다고 가정하면, 아래 가능성들이 우선순위 높은 불연결입니다.

1) 가장 의심되는 원인: 데이터 누수 (Data Leakage)

(1) train/wal split 단위 문제 (예후 훈련)
ST-WL 계층에서 특정 종이 디자인은 케이스입니다.

- split이 sample(case) 단위가 아니라
- spot 단위
- patch 단위
- tile 단위
- 로이어 있으므로

같은 patient / 같은 date / 같은 tissue가 train과 wal에 동시에 들어갑니다

이 경우

1) 가장 유력: 일회값에 NaN / Inf가 섞여 있음

왜 저기 터졌나?

- 샘플 수를 늘리면서
 - 이전엔 없던 저작질 ST 샘플
 - normalization 안 된 sample
 - HVG가 거의 0인 sample
- 이 들어왔을 가능성이 큽니다.

특히 위험한 지점

- ST expression preprocessing:
 - log1p 전에 음수

+ Ask anything

Epoch 3/10
Training: 84% 51/61 [02:59<00:28, 2.90s/it, acc=43.1%, loss=nan]C

이런 식으로 loss=nan이 뜨고 그 뒤로도 계속 유지돼. 예상 원인 알려 줘.

으케이, 이 로그는 원인이 꽤 명확한 NaN 패턴입니다.
'샘플 누리자마자 NaN' - 데이터 분포 + 수치 안정성 문제 쪽 먼저 봐야 합니다.
아래를 발생 가능성 높은 순서대로 정리하겠습니다.



- 모델 학습 과정에서 발생한 과도한 성능 및 NaN loss 문제에 대해, AI를 활용하여 로그 패턴을 분석하고 데이터 누수, 수치 불안정성, AMP(autocast) 사용에 따른 영향 등 가능한 원인을 점검함.
- Autocast(AMP) 사용 여부가 ST-only 모델에서 수치 안정성과 최적화 경로에 미치는 영향을 분석하고, 실험 공정성과 재현성을 고려하여 학습 설정을 통일하거나 예외를 명시적으로 정당화하는 방향으로 실험 설계를 수정함.
- 수정된 실험 설정에서 얻은 학습 로그(accuracy, AUC, precision/recall 변화) 해석을 요청함.

5) 실험 결과 분석

- ChatGPT: <https://chatgpt.com/share/697d5a5c-de3c-8007-a44c-6d9c8ae6a8a9>
- ChatGPT: <https://chatgpt.com/share/697d5a60-66cc-8007-aeb7-974ed0970090>



- Unimodal vs. Multimodal 성능 비교를 위해, 모델별 학습 로그 기반 학습 양상 및 성능 분석을 요청함.
- Epoch별 accuracy, ROC/AUC, precision, recall, f1-score 변화를 통해 학습이 정상적으로 이루어졌는지, 모델이 label 예측에서 어떤 강점과 약점을 가지는지 즉시 확인할 수 있었으며, 긴 출력 로그를 수작업으로 확인하고 분석하는 데 드는 시간을 극단적으로 단축할 수 있다는 점에 AI를 통한 결과 분석의 의의가 존재함.
- Test 출력인 embedding UMAP, scatter plot, gene importance 등을 파악하고, 이를 바탕으로 논문 초안 작성을 요청함. 모델 출력이 여러 파일에 걸쳐 산발적으로 존재하는 상황에서, 이를 한 번에 취합해 결론을 도출함으로써 효율적인 연구를 진행할 수 있음.



7. 논문 작성

1) Abstract 초안 및 논문 개요 제안

- ChatGPT: <https://chatgpt.com/share/695b6108-ffe4-8007-a99b-08f077328270>

The image shows two side-by-side ChatGPT interface windows. The left window displays the abstract section, which includes a detailed description of the model architecture (ST-Informer + fusion layer) and its performance (XAI). The right window displays the discussion section, which includes a brief summary of the model's features and its potential applications.

- 본격적인 논문 작성 전, 논문의 흐름을 파악하기 위해 abstract와 개요 작성은 요청함.
- 이전 대화들로 인해 모델이 본 연구의 목표와 진행 상황을 알고 있는 상황에서 별도의 상세 설명 없이 빠르게 작업할 수 있었으며, 이후 상호 피드백을 통해 원하는 구조의 개요와 설명 흐름을 도출함.

2) Introduction, related work 개요 제안 및 퇴고

- Perplexity: https://www.perplexity.ai/search/daehoe-jeculyong-nonmuneul-jag-u_8rk6MR.OUhAe8bbW78Q#4

The image shows two side-by-side Perplexity interface windows. The left window displays the introduction section, which includes a detailed description of the model's architecture and its performance (XAI). The right window displays the related work section, which includes a brief summary of the model's features and its potential applications.

- Introduction-related work의 개요 및 본문 작성은 요청하고, 상호 피드백을 반복하며 내용을 개선함.

2) Experiment 국문 초안 퇴고 및 번역



- ChatGPT: <https://chatgpt.com/share/697d592d-c950-8007-a14e-9311ea335ff7>
- ChatGPT: <https://chatgpt.com/share/697d58a2-7b14-8007-ad91-b5a680a3e035>
- ChatGPT: <https://chatgpt.com/share/697d5a60-66cc-8007-aeb7-974ed0970090>

4.2 Model Variants Comparison (Ver1-Ver2)

본 장에서는 동일한 텍스트를 하에서 서 가지 보이 변형(Ver1, Ver2)의 예측 성능을 비교하여, 터키어
체 설계 선택에 보이 성능에 차이는 영향을 분석한다. 각 모델은 세 가지 서로 다른 kaggle 전처리방법:
labeled concatenation based, similarity based, gated fusion을 적용하여 허가되었습니다. 이를 통해 교육 구조와 kaggle 허가 간의 상호작용을 함께 관찰하였다.

여기 Ver1baseline 모델은 모두 전처리으로 허가적인 성능을 보였으며, 특히
concatenation 기반 fusion에서 비교적 높은 성능을 기록하였다. 이는 많은 미지수-공전자 표현 활용으로
포함 다른 질병 분야 고백에서 유전학적 패턴의 기능성을 시사하며, 본 연구에서 설정한 baseline의 적합도
를 재정명하였다.

반면, Ver1/similarity encoder 모델은 모든 kaggle 전처리 $\frac{1}{4}$, Ver1 대비 활용되며 낮은 성능을 보였다. 특히
gated fusion 기반 fusion에서 저수준 성능을 보였고, 이는 미지수-공전자 표현 활용 활용으로 인해
전처리 방식에 따른 성능 차이가 커지면서 결과를 저하시켰다.

본 장에서는 차 가지 보이 변형인 baseline + 4 model(Ver1, triple encoder 구조 Ver1L, 그리고
+ residual-encoder)을 활용하여 성능을 평가하였다. 이후 본 장 전반에서는 이를 각각 Ver1, Ver1L, Ver1L로 자
세히 살펴보았다.

- 내용

3) 레퍼런스 검색

- Perplexity: <https://www.perplexity.ai/search/figure-4-img-st-modeleseo-sang-JAHczkG2R4ehgpICbRzQ6A#0>

Figure 4. Img-ST 모델에서 상위 30개의 중요 유전자(top-30)를 기준으로, 동일한 유전자 집합에 대해 ST-only와 Image+BT, 철학에서의 중요도 점수를 비교한 결과를 보여준다.
Image+BT 분류기는 C8,Ctgf, Choc,Lyc2,Cldn20등을 먼저 언급 및 대체로 표기된 경과, 전문적인 표현과 함께, Atp6v0d, O01162, Corf1, Chcr1a등을

Now let me create a final comprehensive report with proper citation:

유전자 액셀 레퍼런스 보고서

Rapport Transcription: 문학의 주요 유전체들과 고민을 학회지 차로

I. 연구 분야 및 보제 경로 유전체

C3 (Complement Component 3)

C3는 보체 캐스터의 핵심 활성화 분분으로서 종양 미세환경에서 면역 체계적 역할을 수행합니다. 종양 유래 CD45 대비 신호전달을 통해 종양-면역 대체세포에서 면역 체계 활성을 증강시키고, C3b+T 세포의 절연과 기능을 억제합니다. 특히 긴급 악력에서 C3b/C3a/C3aR1 복현전 면역 체계 특성과 함께 C3b/C3a/C3aR1 억제합니다.

Ask a follow-up

Ask a follow-up

- Perplexity deep research 기능을 사용하여, ChatGPT가 작성한 gene important 분석에 포함된 gene 특성 설명에 대한 레퍼런스를 검색함. 인용 수나 기관 공신력에 따른 필터링을 요청함으로써, 신뢰도 없는 자료를 출력하거나 hallucination이 발생하는 등의 문제를 사전 예방하고자 함.
- 도메인 지식이 부족하거나 해당 분야의 학회 및 저널에 익숙하지 않은 입장에서, AI를 활용해 레퍼런스 검색을 자동화하여 논문을 빠르게 작성할 수 있었음.

4) Discussion 작성

- ChatGPT: <https://chatgpt.com/share/697dba1f-373c-8007-b17f-06cb9ea54924>



좋은 포인트 정리해 두었습니다. 말씀해 주신 다섯 가지를 과도한 claim 없이, 지금 논문 흐름("baseline-level but meaningful")에 맞춰서
Discussion → Conclusion 구조로 국문+영문 각각 작성해 드리겠습니다.
("바이오 분야 AI agent" 표현은 아래하고 하셔서, 직접적인 agent 주장은 피하고 "agent-like analysis pipeline / foundation" 정도로 완화했습니다.)

5. Discussion (국문)
본 연구는 공간 전사체(ST)와 바이오 이미지(WSI)를 결합한 분티모달 종양 예측 모델을 제안하고, 비교적 제한된 데이터셋과 단순한 모델 구조에서도 의미 있는 성능과 설명 가능성을 확보할 수 있음을 보였다. 특히 ST-WSI 퀘어 데이터가 충분하지 않은 현실적인 연구 노경에서도 얼티모달 접근이 실질적인 이점을 제공할 수 있음을 실험적으로 확인하였다. 첫에서 이어가 있다.

+ Ask anything

most existing histology-transcriptomics models focus on gene- or cell-type-level prediction, with limited emphasis on tumor-level clinical outcomes
related work에서 이거 두개 합치는거 어때

네, 두 단락을 합치는 게 더 좋습니다.
지금 상태에서는 내용이 같지만, reviewer 입장에서는 “그래서 이 둘의 공통 한계가 뭔데?”라는 질문이 자연스럽게 생깁니다.
하나의 흐름으로 묶어서 ‘히스토리 → 공통 한계 → 우리 포지션’이 드러나게 만드는 게 이상적입니다.
아래에 **합친 버전(다들은 문장)**과, 왜 이렇게 합치는 게 좋은지를 함께 정리해 드릴게요.

• 통합 권장 버전 (2.2 Multimodal Integration)

Histology-Genomics and Histology-Transcriptomics Integration
Early multimodal studies integrated histopathology images with genomic data such as somatic mutations, copy number variations, and bulk RNA sequencing, showing that combining morphological and molecular information can improve predictive performance over unimodal models (22, 20). However, these approaches typically rely on simple feature concatenation, treating imaging and genomic modalities as independent channels and limiting their ability to capture structured morphology-molecular relationships (23).

More recent work has extended this paradigm to histology-transcriptomics integration, particularly leveraging spatial transcriptomics as supervision for topography-based models. These studies demonstrate that gene expression patterns can be partially inferred from WSIs and that cross-modal

+ Ask anything

- 실험 내용까지 논문 작성을 완료한 상황에서, introduction부터 experiment까지의 내용을 전체 검토하고 이에 기반하여 discussion, limitation, 그리고 conclusion 작성을 요청함.
- 이후 분량 조절을 위한 내용 조정을 지시하였으며, 최종 퇴고까지 상호 피드백을 반복하며 논문 작성을 완료함.
- 수작업으로 검토할 경우 오랜 시간이 소요될 뿐 아니라 미처 발견하지 못한 오류가 남아 있을 수 있는데, AI를 활용해 빠르고 정확한 검토가 가능함.



8. 자체 리뷰 및 수정

1) 논문 최종본 review

- ChatGPT: <https://chatgpt.com/share/697dba1f-373c-8007-b17f-06cb9ea54924>

*7(4)에서 활용한 링크이나, 이전 대화 history 활용을 위해 해당 채팅에서 최종본 review까지 진행하였으므로 다시 첨부

● 심사 주안점

항목	설명	점수
제작 기준	• 연구가 실제로 유의미하고 실질적인 문제를 다룬↓ 있으며, 사회적 학문적 가치가 있는지	20
제작 내용	• 연구가 학제적 관점으로 접근하여 문제를 주고 있는지, 사회학적 관점으로 접근하여 문제를 주고 있는지	20
제작 방법	• 연구가 정량적·정성적·설증적·비설증적 등 다양한 방법으로 문제를 조망하는지, 그리고 그 결과를 통해 문제를 해결하는지	20
제작 결과	• 연구가 문제를 해결하는 데 있어 학제적 관점으로 접근하여 문제를 주고 있는지, 그리고 그 결과를 통해 문제를 해결하는지	20
인간·사회	• 연구 결과가 개인이나 공동체나 사회·문화·언어·인구·구성인구에게 긍정적 영향을 미친다.	20
제작 특장점	• 연구가 다른 학제적 관점으로 접근하여 문제를 주고 있는지	20
제작 단점	• 연구가 문제를 해결하는 데 있어 학제적 관점으로 접근하여 문제를 주고 있는지	20
제작 저작자	• 저작자는 학제적 관점으로 접근하여 문제를 주고 있는지	20

채점 기준인데, 우리 논문이 여기에 fit한지도 검토해 줘. 항목별로 조건을 만족하는지 아닌지 따져 봐.

좋은 타이밍에 접검하십니다.
이 기준으로 보면 우리 논문은 전반적으로 잘 fit하지만, 몇 항목은 “의도는 충족, 표현은 보완 필요” 상태입니다.
아래에서 **항목별로 ‘충족’ 여부 / 균거 / 리스크 / 보완 포인트**를 넣정하게 나눠서 보겠습니다.

■ 주제의 실용성 (20점) → 충족 (◎)

평가 기준

| 연구가 실제로 유의미하고 실질적인 문제를 다룬↓ 있으며, 사회적 학문적 가치가 있는지

우리 노력에

● 심사 주안점

항목	설명	점수
제작 기준	• 연구가 실제로 유의미하고 실질적인 문제를 다룬↓ 있으며, 사회학적 관점으로 접근하여 문제를 주고 있는지	20
제작 내용	• 연구가 정량적·정성적·설증적·비설증적 등 다양한 방법으로 문제를 조망하는지, 그리고 그 결과를 통해 문제를 해결하는지	20
제작 방법	• 연구가 학제적 관점으로 접근하여 문제를 주고 있는지, 그리고 그 결과를 통해 문제를 해결하는지	20
제작 결과	• 연구가 문제를 해결하는 데 있어 학제적 관점으로 접근하여 문제를 주고 있는지, 그리고 그 결과를 통해 문제를 해결하는지	20
인간·사회	• 연구 결과가 개인이나 공동체나 사회·문화·언어·인구·구성인구에게 긍정적 영향을 미친다.	20
제작 특장점	• 연구가 다른 학제적 관점으로 접근하여 문제를 주고 있는지	20
제작 단점	• 연구가 문제를 해결하는 데 있어 학제적 관점으로 접근하여 문제를 주고 있는지	20
제작 저작자	• 저작자는 학제적 관점으로 접근하여 문제를 주고 있는지	20

채점 기준인데, 우리 논문이 여기에 fit한지도 검토해 줘. “심사자 입장”에서 “항목별로 조건을 만족하는지 아닌지 따져 봐. 인간 심사자, AI 심사자 둘 다로서 한번 해 봐.”

좋습니다. 이번에는 완전히 “심사자 입장”으로 보겠습니다.
의도·노력·역량은 대체로 배제하고, 제출된 논문 텍스트만 보고 판단하는 상황을 가정하겠습니다.
그리고 요청하신 대로 ① 인간 심사자 / ② AI 심사자 두 관점에서 항목별로 합격/부분/미흡을 넣정하게 나누겠습니다.

① 인간 심사자 관점 평가

전체:

- *AI/ML 전공 교수 또는 박사급 연구자*

- 작성한 논문 전체 파일을 전송하여 전체적인 교정을 지시하였으며, 해당 과정 전반에서 논리 전개 흐름과 주장의 일관성 등 주요한 부분부터 비문/오탈자 교정 등의 사소한 요소까지 전체적인 검토를 요청함.
- 공지사항의 “심사 주안점” 표를 첨부해 결과물이 심사 조건에 부합하는지 항목별로 분석하도록 지시하였으며, 특히 심사자 입장에서 판단할 것을 요청하여 보다 객관적이고 비판적인 평가가 가능하도록 함.