10th Anniversary

GISC 2026

Global ICT Standards Conference 2026

ICT 표준특허 세미나 (세션 1)

AI 시대의 ICT 표준특허 전략과 디지털 주권 경쟁

AI 시대의 ICT 표준특허 전략: 영상에서 피지컬 AI까지

이장훈 대표변리사 | 특허법률사무소 베젤

ICT Standards and Intellectual Property:
Shaping the Physical AI

주최 과학기술정보통신부 지식재산처 주관 국립전파연구원 TTA IITP KEA kista ETRI

Index

발표 순서

01 **HEVC 표준특허 구조**
규격이 있어서 계산되는 위험

02 **AI와 영상 표준의 변화**
코덱 안으로 들어오는 신경망

03 **피지컬 AI 영상 표준**
공식표준이 덮는 구간

04 **트랜스포머 사실표준**
트랜스포머 구글 특허

05 **정리와 대응**
공식표준 SEP와 사실표준 특허의 비교

Abstract

발표 개요

ICT 표준특허

피지컬 AI 확산으로 영상 코덱, 무선통신 등 ICT 표준의 적용 기기가 로봇과 자율주행으로 넓어지면서 표준특허의 라이선스와 분쟁 구조도 변화하고 있습니다. **H.265 코덱 표준특허 필수성 검증 사례**를 바탕으로 클레임차트 작성 방법과 판단 기준을 소개하고, 영상 표준 안으로 들어오는 신경망과 피지컬 AI 영상의 특징을 거쳐, 그 끝에 있는 **사실표준(트랜스포머 아키텍처) 특허**의 구조와 국내 기업의 대응 방향을 제시합니다.

01. 영상 파이프라인 위에서 본 표준특허의 위치 변화

HEVC까지는 코덱 구간의 권리자를 알고 있었음. 피지컬 AI에서 가치가 결정되는 구간은 코덱 뒤에 위치

공식표준 구간 (HEVC, VVC, VCM/FCM)

사실표준 구간 (트랜스포머 구조)

카메라·센서 → 코덱 / 특징 압축 → 비전 인코더 (ViT) → 트랜스포머 본체 (VLA) → 행동 출력



출처: Figure AI, “Helix” (2025.02), figure.ai

질문

- 5G, HEVC의 표준특허는 권리자 목록과 라이선스 조건이 공개
- LLM, VLA의 기반 구조에 대한 특허는 누가 보유하는가
- 그 권리에는 필수성 선언, FRAND, 특허풀이 존재하는가

발표 순서

I. HEVC 표준특허 구조: 규격이 있어서 계산되는 위험

II. AI와 영상 표준의 변화: 코덱 안으로 들어오는 신경망

III. 피지컬 AI 영상의 특징과 표준: 공식표준이 덮는 구간

IV. 사실표준: 트랜스포머 구글 특허

V. 정리와 대응

01. 공식표준 체계의 권리 구조: 규격, 선언, FRAND, 특허풀

ITU-T H.265 / ISO/IEC 23008-2 (HEVC) 기준

1 규격 확정
ITU-T VCEG, ISO/IEC MPEG 공동(JCT-VC)으로 syntax, 복호 절차, 프로파일·레벨 확정

→

2 필수성 선언
참여 기업이 특허정책에 따라 규격 실시에 필수인 특허를 선언(declaration)

→

3 FRAND 확약
공정·합리·비차별 조건의 라이선스 제공 의무 수용

→

4 특허풀 라이선스
Access Advance, Via LA 등이 선언 특허를 필수성 평가 후 집단 라이선스

구조 요소	공개되는 정보	기업이 얻는 것
필수성 선언 목록	권리자, 특허번호, 대응 규격 조항	위험의 범위 특정
FRAND 확약	라이선스 거절 불가, 차별 금지	협상 지위 확보
특허풀 요율표	제품 단위 요율, 상한, 지역 구분	비용 예측 가능
분쟁 해결 경로	필수성 다툼, 무효심판, 요율 소송 선례	대응 절차 정형화

01. 필수성 검증 실무: 청구항과 규격 조항의 클레임차트 대응

H.265 선언 특허의 클레임차트 예시

클레임차트 구조 (US 9300962 B2, 타일 병렬 복호)

청구항 구성요소	H.265 규격 대응 조항	판정
전제부: 비디오 복호기, 타일 병렬 처리	§6.1 비트스트림 형식, §8.1.1 복호 절차	대응
구성 A: 타일 사용 여부를 나타내는 제1 신택스 요소	§7.3.2.3.1 PPS syntax, §7.4.3.3.1 tiles_enabled_flag	문언 대응
구성 B: 타일 진입점 위치 정보	§7.3.6.1 slice segment header, §7.4.7.1 entry_point_offset	문언 대응
구성 C: 타일 단위 CTB 주사 순서	§6.5.1 CTB raster and tile scanning, §3.174, §3.177	대응
구성 D: 타일별 엔트로피 복호 초기화	§9.3.2.1 initialization process, §7.3.8.1	대응

판정 체계: 문언 대응 / 조건부 필수(플래그·프로파일 의존) / 비매칭(규격 선택 사항, 구성요소 부재)

검증이 가능한 이유

규격 문서가 복호기의 필수 동작을 조항 단위로 고정하므로, 청구항 구성요소마다 대응 조항의 존재 여부를 문언으로 판정 가능. **권리의 범위와 가치가 규격 문서 안에서 결정되는 구조**

01. 필수성 검증 실무: 검증 절차와 판정 기준

청구항 구성요소 단위 분해와 규격 조항 매핑, 세 가지 판정 유형

검증 절차

1 대상 특정
선언 DB에서 특허와 대응 규격 버전 확인

2 프로파일 확정
제품이 실시하는 프로파일·레벨·툴 범위 확정

3 구성요소 분해
청구항을 전제부와 구성요소 단위로 분해

4 조항 매핑
구성요소별 syntax, semantics, 복호 절차 조항 대응

5 판정
전 구성요소 대응 시 필수, 하나라도 부재 시 비매칭

판정 유형

문언 대응

필수

규격의 필수 동작이 구성요소를 그대로 기술. 규격을 따르는 모든 복호기가 해당 동작을 실시하므로 규격 준수만으로 충족

조건부 필수

프로파일 확인

플래그 값이나 프로파일에 따라 동작이 달라지는 툴. 해당 툴을 사용하는 프로파일·플래그 조건에서만 필수이므로 제품 사양 대조 필요

비매칭

필수 아님

규격이 강제하지 않는 선택 사항(부호화기 측 결정 과정 등)이거나 구성요소가 규격에 부재. 규격 준수와 무관하게 회피 가능

2026년 국내 검증 사례 V사에서 소송 목적으로 양수도된 H.265 선언 특허 대상: **필수 84%**(Main 80%, REXt 4%), 일부 청구항 비매칭 4%, 비매칭 12%

01. 공식표준 체계가 실시 기업에 주는 세 가지

권리자 목록, 요율, 분쟁 경로의 사전 공개

1 위험의 가시성

- 선언 특허 DB로 권리자, 건수, 대응 조항 파악
- 제품 프로파일 기준 필수 청구항 범위 산정
- 라이선스 전 자체 필수성 검증 수행 가능

2 협상 구조의 정형화

- 풀 요율표와 상한으로 비용 예측
- FRAND 확약으로 라이선스 거절, 차별 조건 제한
- 필수성 다툼, 무효심판, 요율 소송의 선례 축적

3 NPE도 구조 안에서 행사

- 오디오·영상 NPE의 매입 대상은 선언 특허, 풀 이탈 특허
- 행사 시점에도 FRAND 선언의 효력 유지
- 대응 정보 제공 사업의 분석 단위가 규격 조항

I장 정리

HEVC 체계에서 표준특허 위험은 **규격 조항 단위로 특정**되고, **요율표로 계산**되며, **정형화된 경로로 협상** 가능. 이후 장에서 이 세 가지가 AI 영상 파이프라인에서 어디까지 유지되는지 검토

02. 코덱 규격 내부로 진입하는 신경망

규격이 알고리즘 툴이 아니라 학습된 네트워크 구조를 기술하기 시작

JVET NNVC

ITU-T VCEG, ISO/IEC MPEG 공동 탐색

- 신경망 기반 루프 필터, 인트라 예측, 초해상 탐색
- 참조 소프트웨어에 네트워크 구조·가중치 포함
- 차세대(H.266 이후) 규격 후보 툴

ECM

VVC 이후 탐색 모델

- VVC 대비 부호화 이득 확보 툴 집합
- 일부 툴은 학습 기반 파라미터 사용
- 규격 채택 시 툴 단위 필수성 판단 대상

JPEG AI

ISO/IEC 학습 기반 이미지 코딩 (2025 발간)

- 인코더·디코더 자체가 신경망
- 규격이 네트워크 구조, 텐서 형식, 엔트로피 모델 기술
- 복호기의 필수 동작이 가중치에 의존

표준 번호·단계: 2026년 정합성 분석 과제 자료 기준 기재

표준특허 관점의 변화

규격이 네트워크 구조와 가중치를 참조하면, 필수성 검증 대상이 **syntax element에서 네트워크 구조·학습 절차로 이동**. 청구항과 규격의 대응 관계가 조항 단위에서 구조 단위로 전환

02. 영상 소비 주체의 변화: 사람에서 기계로

MPEG VCM(Video Coding for Machines), FCM(Feature Coding for Machines)

사람이 보는 영상 (HEVC, VVC)

카메라 → 인코더 → 비트스트림 전송 → 디코더 (픽셀 복원) → 디스플레이, 시청자 → 품질 지표 **PSNR, SSIM, 주관 화질**

기계가 읽는 영상, 픽셀 (VCM)

카메라 → 기계 최적화 인코더 → 비트스트림 전송 → 디코더 (픽셀 복원) → 인식 모델 → 품질 지표 **mAP, MOTA 등 인식 성능**

기계가 읽는 영상, 특징 (FCM)

카메라 → 특징 추출 앞단 (백본 분할) → 특징 텐서 압축·전송 → 특징 추출 뒷단 → 인식 모델 → 품질 지표 **mAP, MOTA 등 인식 성능**

☐ 규격이 복호 동작을 고정하는 구간 ☐ 표준화 기구 밖에서 설계된 네트워크 구간

FCM의 구조적 의미

규격은 백본 분할 지점과 특징 텐서 형식을 정하지만, 분할되는 네트워크 자체는 표준화 기구 밖에서 설계된 구조. **공식표준이 사실표준 위에 얹히는 첫 사례**

02. 표준특허 구조에 생기는 이중 구조: 공식표준 규격, 사실표준 기반

규격은 공식표준, 그 아래 네트워크 구조는 표준화 기구 특허정책 밖의 권리

VCM / FCM / JPEG AI 규격
(공식표준: 선언, FRAND, 풀 적용)

특징 추출 백본, 학습 기반 코딩 네트워크
(규격이 참조하는 구조)

트랜스포머 등 기반 아키텍처
(사실표준: 선언 없음, 단일 권리자)

선언 대상

선언 범위 밖

특허정책 적용 범위

항목	HEVC 체계	VCM / FCM 체계
필수성 검증 대상	syntax, 복호 절차	네트워크 구조, 학습 절차
규격이 참조하는 것	툴 알고리즘	백본 네트워크, 텐서 형식
특허정책 적용	규격 전체	규격 문언까지, 백본 구조 제외
권리자 파악	선언 DB	선언 DB + 비선언 아키텍처 특허

II장 정리

규격이 신경망을 참조하는 순간, 실시 기업은 규격 선언 특허와 별도로 **규격이 참조하는 네트워크 구조의 권리를 추가 검토**해야 하는 상태. 그 구조의 권리자는 표준화 기구 특허정책에 구속되지 않음

03. 피지컬 AI 영상의 네 가지 특징

로봇, 자율주행 기기가 소비하는 영상의 조건



출처: Tesla, “AI & Robotics”, tesla.com

1 소비자는 모델

VLM, VLA가 영상을 직접 수신. 사람의 시청 과정 부재. 품질 기준은 행동 성공률

2 폐루프 제어 지연

관측, 판단, 행동이 프레임 단위로 반복. π0 기준 50Hz 행동 생성. 압축·전송 지연이 제어 성능에 직결

3 다중 센서 융합

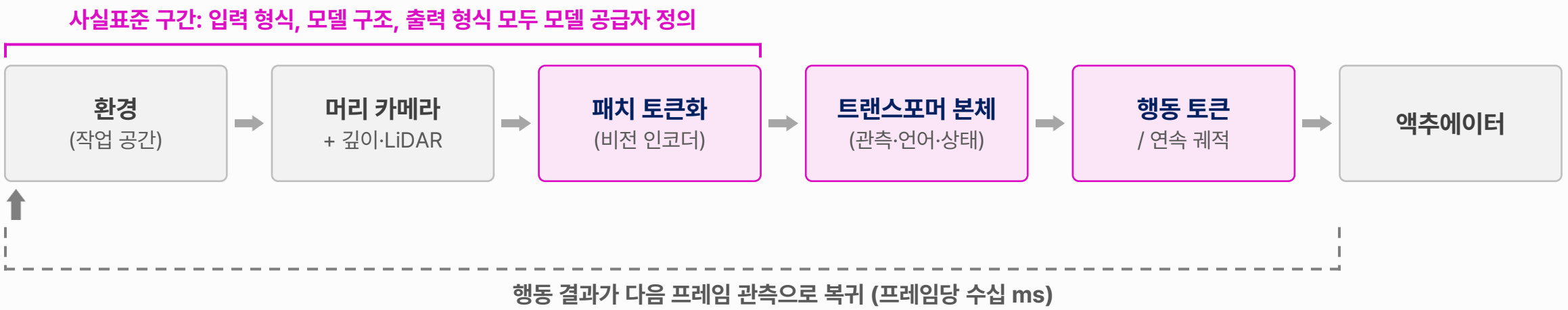
머리 카메라 기본, 손목 카메라 선택 사양. 깊이, LiDAR, 이벤트 카메라 병행. 시점 정합 전제

4 토큰으로 소비

픽셀 복원 없이 패치 임베딩으로 변환. 비전 트랜스포머가 패치 시퀀스를 토큰화

03. 피지컬 AI 페루프: 영상이 소비되는 경로

로봇, 자율주행 기기가 소비하는 영상의 조건



요지

네 가지 특징 모두 영상의 최종 소비 지점을 **트랜스포머 기반 모델로 고정**. 영상 표준의 평가 기준이 화질에서 모델 성능으로 이동

03. 피지컬 AI 영상 파이프라인의 계층별 표준 현황

압축, 인터페이스 계층은 공식표준, 모델 입출력 계층은 사실표준

계층	기능	공식표준	사실표준
행동 출력	행동 토큰, 연속 궤적	부재	모델 공급자별 토큰화 방식
모델 본체	관측에서 행동으로 변환	부재	트랜스포머 아키텍처 (단일 권리자)
토큰화·모델 입력	패치 임베딩, 해상도, 정규화	부재	비전 인코더별 전처리 규약
기계 소비 압축	엣지·클라우드 분산 추론	VCM, FCM (진행)	
점군·3D	LiDAR, 깊이 데이터	G-PCC, V-PCC	
전송·저장 압축	원격 관제, 로그 저장	HEVC, VVC	
센서 인터페이스	카메라, LiDAR 출력 형식	MIPI CSI, 차량용 이더넷	로봇 미들웨어 메시지

☐ 모델 입출력 계층 (사실표준) ☐ 압축, 인터페이스 계층 (공식표준)

표준 번호·단계: 2026년 정합성 분석 과제 자료 기준 기재

구조

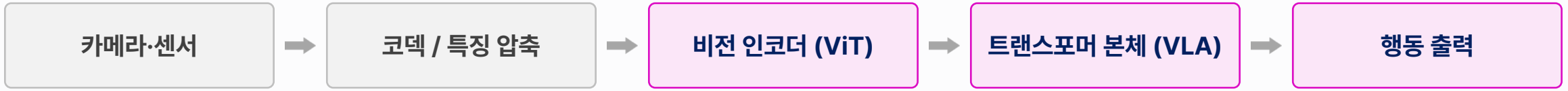
공식표준이 덮는 계층은 모델 입력 직전까지. 위 세 계층은 모두 모델 공급자가 정의한 사실표준이며, **그 공통 기반이 트랜스포머**

03. 파이프라인 위 표준 적용 범위와 가치 결정 구간의 불일치

공식표준 구간은 앞쪽 두 단계, 차별화와 가치는 비전 인코더 이후 구간에서 결정

공식표준 적용: 선언, FRAND, 풀

사실표준: 선언 없음, 단일 권리자, 라이선스 재량



코덱 / 특징 압축

HEVC, VVC, VCM, FCM. 권리자 목록과 요율 공개. 규격 준수 시 회피 불가하나 비용 계산 가능

비전 인코더

ViT 계열(SigLIP, DINOv2 등). self-attention 인코더 구조. 원천특허 청구항 1의 인코더 구성과 대응 가능

트랜스포머 본체

Llama, Gemma 계열 디코더. 이미지 토큰과 텍스트를 받아 행동 토큰 생성. Decoder-only 특허와 대응 가능

III장 정리

피지컬 AI에서 영상의 가치는 코덱이 아니라 **인코더와 본체에서 결정**. 그 구간의 공통 구조에 대한 권리는 공식표준 체계 밖에 위치. IV장에서 그 권리의 실체 확인

04. 트랜스포머 특허, 논문, 모델의 시간 순서

구글 가출원 우선일이 논문 공개와 GPT 발표보다 선행

2017-05-23
구글 가출원
US 62/510,256
Encoder–Decoder 계열 우선일

2017-10-27
구글 가출원
US 62/578,358
Decoder-only 계열 우선일

2019-10-22
미국 등록
US10452978B2
Encoder–Decoder 원천특허

2024-01-08
계속출원
US18/407,299
원천 계열 청구항 재설계



2017-06-12
논문 공개
Attention Is All You Need
우선일 20일 후 arXiv 공개

2018-06-11
OpenAI
GPT 최초 발표
Decoder-only 우선일 이후

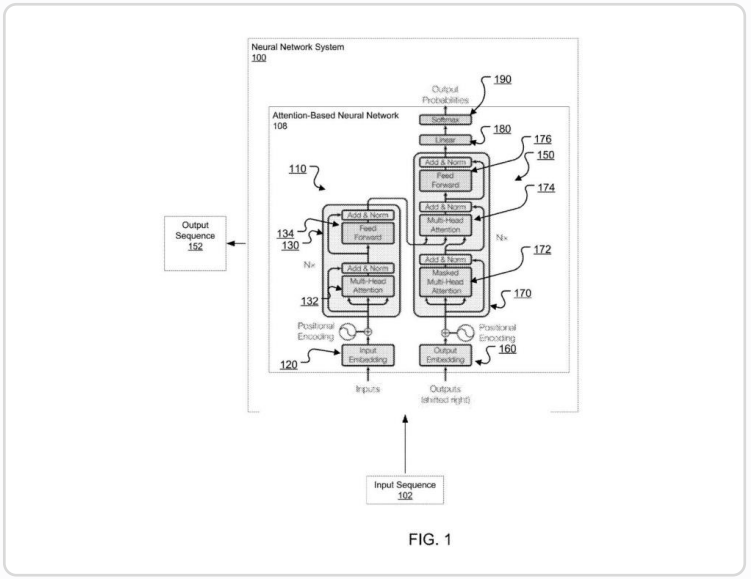
2023-01-09 / 01-17
한국·미국 등록
KR102486348B1 · US11556786B2
원천 KR 등록, Decoder-only US 등록

2025
미국 등록
US12217173B2
원천 계열 후속 등록, 계류 2건 추가

● 특허 사건 ● 논문·모델 발표 우선권 주장일과 각 청구항의 실질 유효출원일은 별도 확인 대상

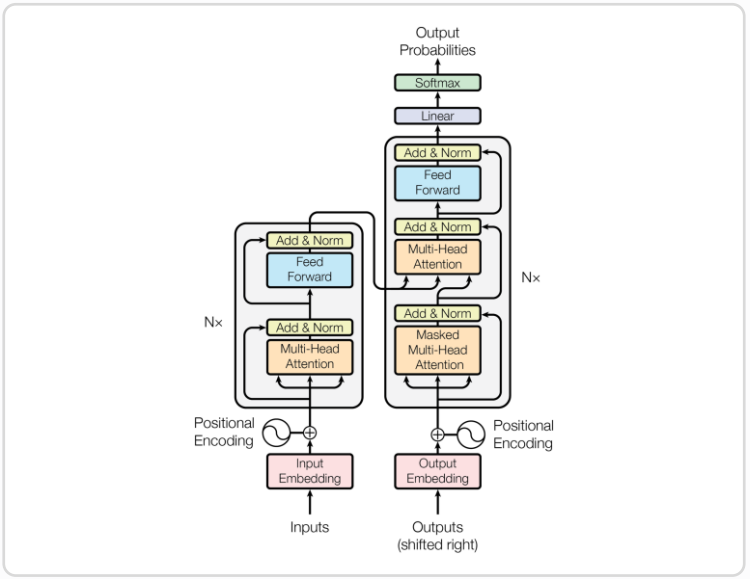
04. 두 특허의 대표도와 논문 Figure 1

Encoder–Decoder 구조와 Decoder-only 구조. 공보 FIG. 1 원문 (USPTO 등록공보, Google Patents 제공 PDF)



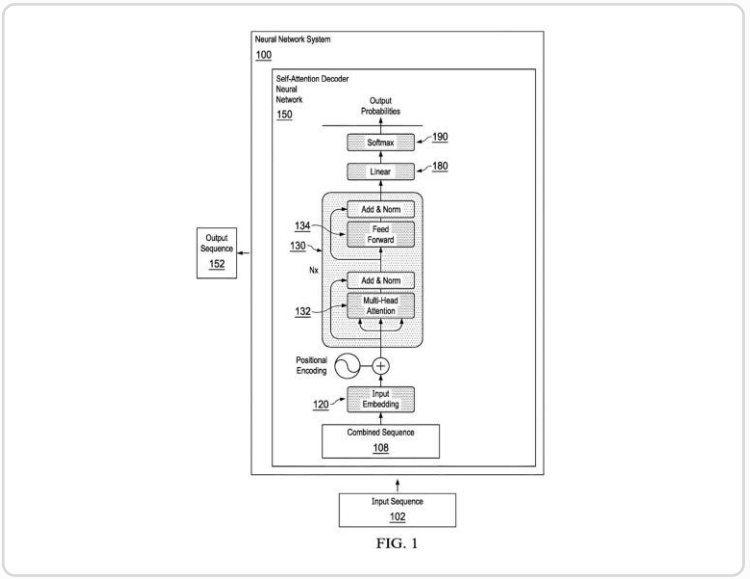
US10452978B2 FIG. 1: 인코더(110)와 디코더(150)

2017-05-23 우선일



Vaswani et al. (2017) Figure 1, arXiv

2017-06-12 공개



US11556786B2 FIG. 1: 디코더(150) 단일

2017-10-27 우선일

논문 Figure 1의 인코더 블록(좌)과 디코더 블록(우)이 US10452978B2 FIG. 1의 인코더(110)·디코더(150)에 대응. US11556786B2 FIG. 1은 그중 디코더 블록만 보유. **가출원일은 논문 공개 20일 전**

04. 두 구조의 차이: Encoder–Decoder와 Decoder-only

US10452978B2 (2017-05-23 우선일)와 US11556786B2 (2017-10-27 우선일)의 대비

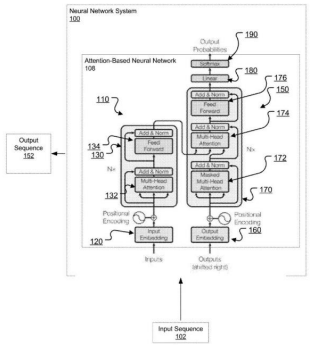


FIG. 1

US10452978B2

구분	US10452978B2 (Encoder–Decoder)	US11556786B2 (Decoder-only)
인코더	self-attention 인코더(110) 존재	부재
디코더 입력	인코더 출력 + 이전 출력(shifted right)	입력 토큰 + 기생성 토큰의 combined sequence(108)
디코더 어텐션	마스크드 셀프어텐션(172) + 인코더-디코더 어텐션(174)	마스크드 셀프어텐션(132)
청구항 1의 요구	인코더와 그 출력을 받는 디코더	combined sequence, masked self-attention, score distribution, 토큰 선택
대조 대상	비전 인코더 결합 VLM/VLA	자기회귀 행동 토큰 생성 모델

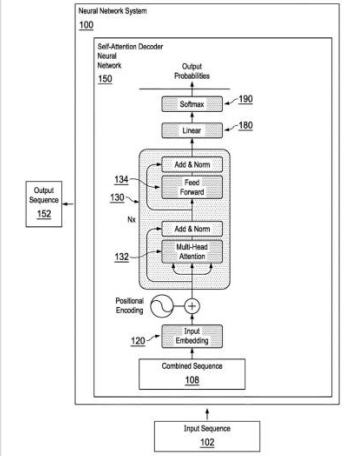
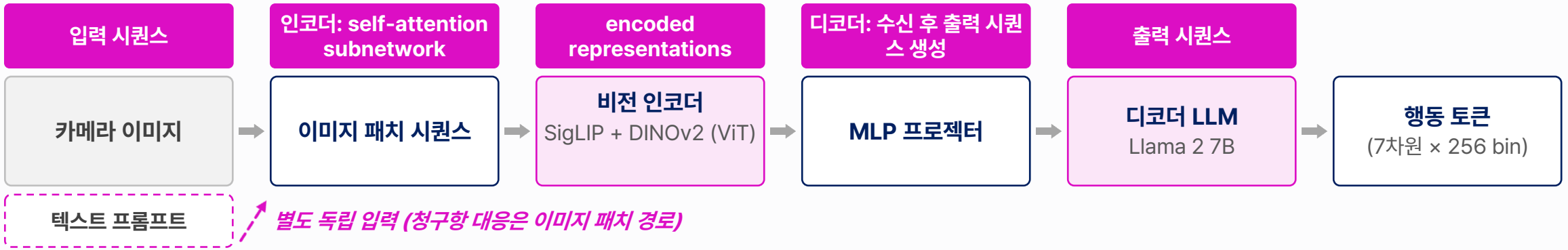


FIG. 1

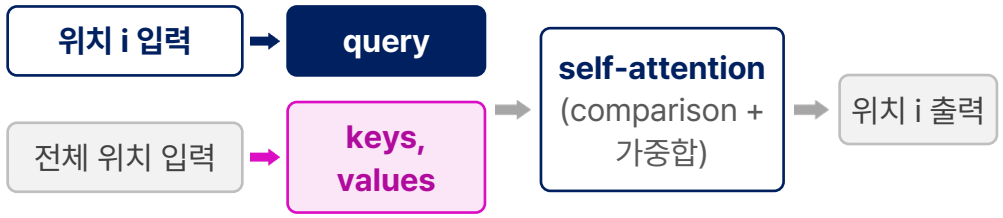
US11556786B2

04. VLA 구조 위에 놓은 원천특허 청구항 1의 구성요소

OpenVLA 공개 구조 기준. 청구항 구성요소를 강조색 태그로 표시



인코더 subnetwork 내부 (청구항 1의 Q·K·V 한정)

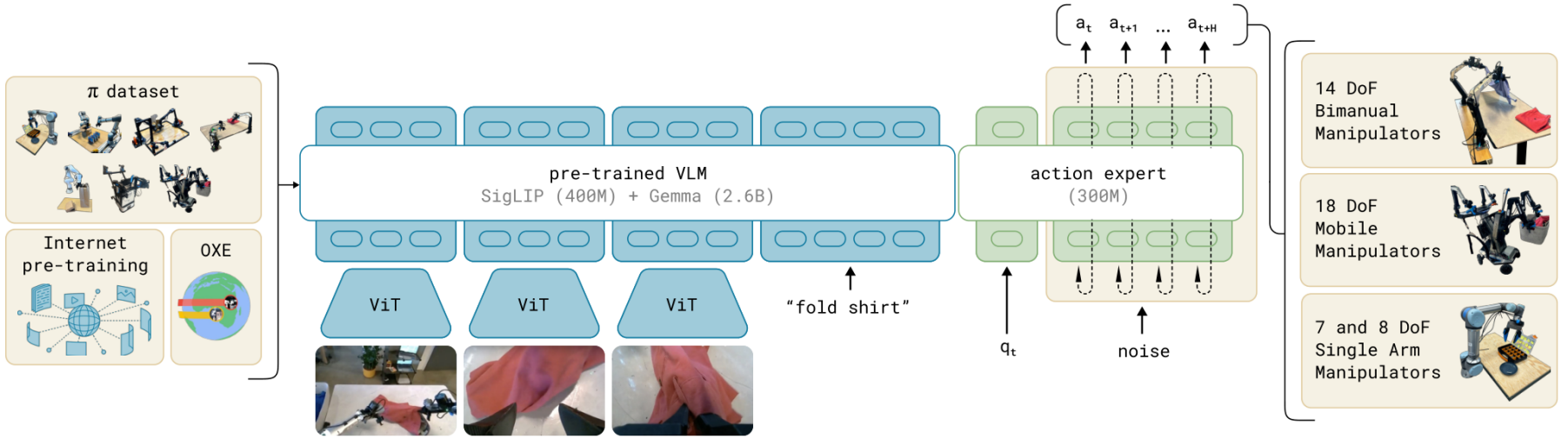


구성요소별 대응 (공개 구조 기준)

- 입력 시퀀스, 인코더, Q·K·V, encoded representation: ViT 구조와 문언 대응
- 디코더의 수신: 프리픽스 토큰 입력. 청구항 1은 수신 방식에 침묵 (크로스어텐션은 종속항 19~23)
- 텍스트 프롬프트는 별개 독립 입력, 이미지 패치 경로로 대응
- 프로젝터 개재: 비침해 논거로서의 효력은 해석 쟁점

04. 공개 논문 구조 위의 대응: π0 아키텍처

Physical Intelligence π0 논문 Figure 3: 비전 인코더(ViT), 트랜스포머 본체(Gemma), 행동 전문가(action expert)의 배치



출처: Black et al., "π0: A Vision-Language-Action Flow Model for General Robot Control" (2024.10), arXiv, Figure 3

입력 시퀀스, 인코더

카메라 이미지(복수 시점)가 SigLIP ViT에서 패치 토큰으로 변환. 청구항 1의 self-attention 인코더 구성에 대응

디코더의 수신

Gemma 2.6B 본체가 이미지 토큰과 언어 지시("fold shirt")를 하나의 시퀀스로 수신. 인코더 출력 수신에 대응

행동 출력의 형식

action expert(300M)가 연속 행동 a_t~a_t+H를 flow matching으로 생성. 토큰 선택 방식(π0-FAST)과 대조 범위 상이

04. 원천특허 US10452978B2 청구항 1과 VLA 구조의 대응

공개 논문 구조 기준 문언 대조. 침해 판단이 아닌 검토 범위 제시

US10452978B2 청구항 1 (시스템항)	OpenVLA (2024)	π0 / π0-FAST
입력 위치별 네트워크 입력을 갖는 입력 시퀀스	이미지 패치 시퀀스 (patch-as-token)	카메라 이미지 패치 시퀀스
인코더: self-attention sub-layer를 갖는 encoder subnetwork 시퀀스	SigLIP, DINOv2 비전 트랜스포머 (ViT 층이 subnetwork)	SigLIP 비전 인코더 (PaliGemma)
특정 위치 입력에서 query, 전체 위치 입력에서 keys·values 도출	ViT self-attention 연산	ViT self-attention 연산
각 입력의 encoded representation 생성	패치 특징 벡터 (채널 결합)	이미지 토큰 임베딩
디코더: encoded representations를 수신해 출력 시퀀스 생성	Llama 2 7B가 MLP 투영 패치 토큰을 입력으로 수신, 행동 토큰 생성	Gemma 본체가 이미지 토큰을 프리픽스로 수신, 행동 출력

04. 원천특허 US10452978B2 청구항 1의 한정 범위

공개 논문 구조 기준 문언 대조. 침해 판단이 아닌 검토 범위 제시

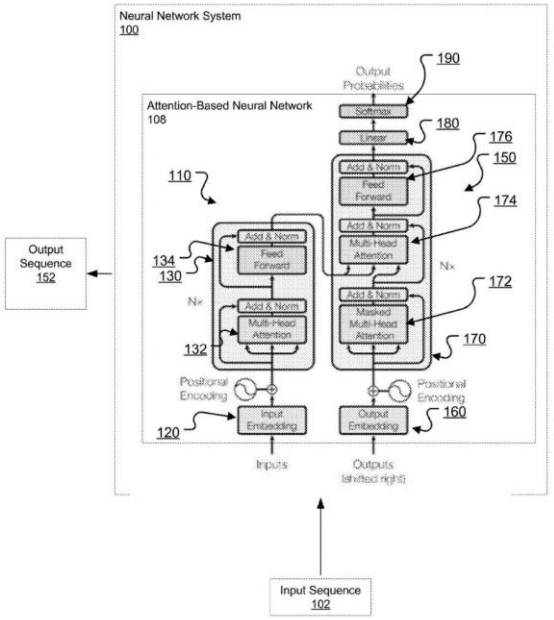


FIG. 1

US10452978B2 FIG. 1: 인코더(110)와 디코더(150)

청구항 1에 없는 한정 (종속항에서 추가)

인코더-디코더 크로스어텐션	청구항 19~23
마스크드 디코더 셀프어텐션	청구항 24~25
멀티헤드	청구항 9, 11, 12
자기회귀 생성	청구항 13
softmax, 스케일링	미기재 (comparison function만 기재)

역설

텍스트 전용 Decoder-only LLM은 인코더 부재로 이 청구항에서 벗어나는 구조. 비전 인코더를 결합한 피지컬 AI 모델은 self-attention 인코더와 그 출력을 받는 디코더를 모두 보유. **영상 입력이 추가되는 순간 가장 넓은 원천 청구항이 대조 대상으로 복귀**

04. 원천특허 대조 시 쟁점 다섯 가지

청구항 해석과 실제 구현 확인이 필요한 지점

1 수신 방식

원 논문·명세서는 크로스어텐션으로 수신, VLA는 프리픽스 토큰으로 수신. 청구항 1 문언은 수신 방식에 침묵(크로스어텐션은 종속항 19~23). 명세서·심사경과 기준 해석 쟁점

2 프로젝터 개재

MLP 투영을 거친 표현이 encoded representations에 해당하는지. 종속항 3(마지막 subnetwork 출력)과의 관계

3 π0 계열 행동 생성

flow matching 디노이징은 토큰 분포를 생성하지 않아 Decoder-only 특허 청구항 1의 score distribution 요소와 불일치 가능. 원천특허는 별개 검토

4 한국 등록 청구항

KR102486348B1 청구항 1은 디코더 대신 제1 신경망, 제2 신경망으로 기재. 디코더 성격 한정이 더 약하므로 국내 서비스 FTO는 한국 청구항 기준 대조 필요

5 권리자의 시장 지위

RT-2 등 구글 자체 VLA 보유. 권리자가 아키텍처 특허 권자이자 피지컬 AI 시장 참여자인 구조

유보

공개 논문 구조 기준 문언 대조이며, 실제 구현, 실시 주체, 관할, 라이선스, 무효 가능성에 따라 결론 상이. 특정 제품의 침해 판단 아님

04. Decoder-only 특허 청구항 1의 생성 루프와 행동 토큰 생성

US11556786B2 청구항 1의 네 단계를 OpenVLA 행동 생성 절차에 배치

① 입력 토큰 + 기생성 출력 토큰의 결합 시퀀스	② masked self-attention decoder 처리	③ 점수 분포 생성	④ 토큰 선택
combined sequence (생성 시점 t) 이미지 토큰 | 텍스트 토큰 | a1 | a2 | … | a(t-1)	**self-attention decoder** (masked self-attention 층 복수)	**출력 토큰 집합의 score distribution**	**다음 행동 토큰 a(t) 선택**

선택된 토큰을 시퀀스 말미에 추가하고 다음 생성 시점으로 반복 (7차원 행동 = 7회 반복)

OpenVLA

행동 7차원을 256 bin으로 이산화, 어휘의 최저빈도 256 토큰을 행동 토큰으로 대체. **①~④ 전부 대응**

π0-FAST

행동을 DCT + BPE로 토큰화해 자기회귀 생성. **①~④ 전부 대응**

π0 (flow matching)

행동을 연속값으로 디노이징 생성. **③ score distribution, ④ 토큰 선택 부재**. 원천특허('978)는 별개 검토

요지

행동을 이산 토큰으로 자기회귀 생성하는 계열은 Decoder-only 청구항의 네 단계와 구성 대응. 별도 인코더의 존재는 청구항 문언상 배제 요소가 아님

04. Decoder-only 특허 청구항 1과 행동 생성 방식의 대응

US11556786B2: 우선일 2017-10-27, 등록 2023-01-17. 후속 US11886998B2

US11556786B2 청구항 1 (방법항)	OpenVLA	π0	π0-FAST
각 생성 시점에 입력 토큰 + 기생성 출력 토큰의 combined sequence 구성	프롬프트 + 생성된 행동 토큰	해당 없음 (프리픽스 비인과, 행동은 연속값)	프롬프트 + 생성된 행동 토큰
복수의 masked self-attention 층을 가진 self-attention decoder로 처리	Llama 2	부분 (액션 전문가 구조)	Gemma
출력 토큰 집합에 대한 score distribution 생성	7차원 × 256 bin 행동 토큰 분포	없음 (flow matching 디노이징)	DCT + BPE 행동 토큰 분포
분포로부터 다음 출력 토큰 선택	자기회귀 next-token 선택	없음	자기회귀 next-token 선택

청구항 구조의 특징

- 별도 인코더 유무에 침묵. 인코더 결합 시스템도 디코더 동작이 문언에 해당하면 대조 대상
- 독립항 1(방법), 12(시스템), 20(저장매체)
- 원 논문(2017-06-12 공개)이 이 계열 우선일(2017-10-27)보다 앞서므로 진보성 판단에서 선행기술 자격 검토 대상

행동 생성 방식에 따른 분기

행동을 이산 토큰으로 자기회귀 생성하는 계열(OpenVLA, π0-FAST, RT-2)은 **Decoder-only 청구항과 구성 대응**. 연속 궤적을 디노이징으로 생성하는 계열(π0)은 score distribution 요소에서 벗어나나, 비전 인코더 결합으로 원천특허 대조는 유지

04. 계속출원 전략: 한 명세서에서 반복되는 독립항 재설계

원천 계열 등록 6건, 계류 2건, 포기 1건. 시장 구조 변화에 맞춘 청구항 재단

원천 계열
2017-05-23 가출원 명세서 (Encoder–Decoder)

- **US10452978** 2019-10 등록
- **US10719764** 2020 등록
- **US2021/0019624** 포기
- **US10956819** 2021 등록
- **US11113602** 2021 등록
- **US11893483** 2024 등록
- **US2024/0144006** 계류
- **US12217173** 2025 등록
- **US2025/0217644** 계류

Decoder-only 계열
2017-10-27 가출원 명세서

- **US2020/0342316** 공개
- **US11556786** 2023-01 등록
- **US11886998** 2024 등록

등록 계류 공개 포기·공개 출원·등록 순 배열. 등록연도는 특허번호 구간 기준 추정, 공식 권리원장 대조 필요

공식표준 특허와의 차이

공식표준 특허는 규격 확정 후 선언되어 범위가 고정. 사실표준 특허는 **명세서 하나로 계류를 유지하며 시장이 채택한 구조에 맞춰 독립항을 계속 재설계 가능**

04. 공식표준 SEP에 있고 트랜스포머 특허에 없는 네 가지

오픈소스 라이선스의 한계와 분쟁 부재의 의미

필수성 선언	FRAND 확약	특허풀	권리자
부재	부재	부재	단일

Apache 2.0 라이선스의 보호 범위

- 제3조 특허 라이선스는 기여물 단독 또는 기여물과 저작물의 결합에 의해 필연적으로 침해되는 청구항에 한정
- 논문 공개는 특허 허락이 아님
- 구글 공개 소프트웨어(Tensor2Tensor 등)의 라이선스가 독자 재구현 모델까지 면책하지 않음
- 특허소송 제기 시 라이선스 종료 조항 별도 확인 대상

분쟁 부재의 의미

- 김동준(2026.08, 산업재산권 제84호): Transformer 특허가 존재하나 분쟁이 나타나지 않은 상황을 OSS 라이선스, 방어적 특허정책, 블랙박스 증명 문제, 실시 주체 귀속 관점에서 분석
- **분쟁 부재는 권리 부재가 아니며, 권리자가 행사 시점을 보류하고 있는 상태**
- PACER, PTAB, ITC, UPC 기록 기준 부존재 증명은 별도 과제

05. 공식표준 SEP와 사실표준 특허의 구조 비교

HEVC 체계와 트랜스포머 특허의 여섯 항목 대비

항목	HEVC SEP (공식표준)	트랜스포머 특허 (사실표준)
권리자	다수, 선언 목록 공개	단일 기업
필수성	선언, 풀 평가, 검증 체계 존재	선언 없음, 청구항 대조로만 확인
라이선스 조건	FRAND, 풀 요율표	권리자 재량
행사 시점	상시 라이선스 협상	권리자 선택
무효 자료	표준 기고문 등 선행기술 풍부	논문 공개일(2017-06-12)과 각 청구항 우선일 대조
설계 회피	규격 준수 시 불가	구조 변경 가능하나 생태계 이탈 비용 발생
청구항 범위 변동	규격 확정 후 고정	계속출원으로 시장 채택 구조에 맞춰 재설계

디지털 주권 관점

국내 독자 파운데이션 모델, 국가 AI 컴퓨팅 인프라, 피지컬 AI 로봇 모두 트랜스포머 구조 위에 구축. 가중치와 데이터의 주권을 확보해도 아키텍처 권리는 외부 단일 기업 소유. **라이선스 요구 발생 시점에 대체 구조로 이전할 수 없는 상태가 위험의 실체**

05. 사실표준 특허의 침해 입증 경로 변화

확산된 아키텍처와 확산되는 역공학 도구: 블랙박스 증명 장벽의 약화

25만+

논문 인용 수 (2026 기준)

Attention Is All You Need (2017)

21세기 최다 인용 논문군. 트랜스포머 구조의 채택 범위이자 원천특허 대조 대상이 될 수 있는 제품·서비스 범위의 지표

10만+

GitHub 스타 수 (2026.10.12)

REA (Reverse Engineer Anything)

AI 코딩 에이전트를 Ghidra 등 역공학 도구에 연결해 앱·바이너리·펌웨어를 소스 코드 없이 분석. 10.6 1만 미만에서 10.12 10만 이상으로 증가

기존 장벽

- 모델 구조가 제품 내부에 있어 실시 사실 확인 곤란 (블랙박스 증명 문제)
- 공개 논문 구조와 실제 구현의 일치 여부 미확인

변화

- 로봇·차량 펌웨어, 온디바이스 모델의 구조 분석 비용 하락
- 청구항 구성요소 단위 확인 가능성 확대

시사점

- 행사 보류 상태를 지탱하던 입증 비용 요인의 약화
- EULA 역공학 금지 조항, 증거능력 쟁점은 병존
- 자사 제품 바이너리 기준 노출 범위 점검 필요

출처: Wikipedia “Attention Is All You Need” (Nature 2025.04 최다 인용 논문 기사 인용), KuCoin News 2026.10.12 REA 보도

05. 계산 구조가 없는 위험에 대한 대응 방향

기업 차원과 기관·정책 차원

기업

- **VLM/VLA 실제 구조 기준 FTO:** 비전 인코더 결합 여부, 프리픽스 수신 방식, 행동 토큰화 방식을 청구항 구성요소와 1:1 대조
- 한국 등록 청구항(제1 신경망, 제2 신경망 문언) 기준 국내 서비스 검토
- 사용 중인 오픈소스 코드의 라이선스 적용 범위 확인(기여물 결합 한정)
- 구조 변형(어텐션 변형, 비 트랜스포머 계열 혼합) 적용 시 설계 근거 기록
- 토큰화, 행동 출력, 센서 융합 계층에서 자사 특허 선점

기관·정책

- **사실표준 특허 모니터링 체계:** 계속출원 공개공보 추적, 한국 패밀리 청구항 관리
- **무효 자료 사전 확보:** 2017-05-23, 2017-10-27 이전 선행기술 정리, 논문 공개일 기준 청구항별 유효 우선일 검토
- VCM/FCM 등 AI 영상 표준화 참여 시 백본 네트워크에 대한 특허정책 적용 범위 문제 제기
- 표준화 기구 거버넌스 표준에 그치지 않고 기술 규격 항목에 국내 특허 반영

결론

공식표준은 특허 위험을 계산 가능하게 만든 체계. 사실표준에는 그 계산 구조가 부재. 피지컬 AI의 영상이 트랜스포머로 수렴하는 지금, **그 구조를 누가 설계하느냐가 디지털 주권 경쟁의 한 면**

10th Anniversary

GISC 2026

Global ICT Standards Conference 2026

감사합니다

이장훈 대표변리사 | 특허법률사무소 베젤

jhlee@vezellaw.com

ICT Standards and Intellectual Property:
Shaping the Physical AI

주최 과학기술정보통신부 지식재산처 주관 국립전파연구원 TTA IITP KEA kista ETRI