

CV+ML 38년의 흐름

4-depth Phylogenetic Taxonomy로 본 paradigm shift

CV+ML Paper Phylogeny project

CVPR · NeurIPS · ICML · ICCV · ICLR · ECCV · 3DV, 1987–2025

April 28, 2026

왜 이런 분류가 필요한가

- CV+ML은 38년 동안 **최소 6번** 패러다임이 통째로 갈렸다.
- 같은 주제도 논문마다 표현이 다름
(예: “image segmentation” \approx “pixel-wise labeling” \approx “dense prediction”)
- 단순 keyword/TF-IDF는 **표현은 다른데 같은 의미인** 클러스터를 못 잡음.
- \Rightarrow **Semantic synonym cluster + 4-depth phylogenetic tree**

이 슬라이드의 목표

8개의 5년 cohort로 잘라서, 각 시기의 **paradigm signature**와 **rank 변화**를 한눈에 보기.
Taxonomy 자체가 시간의 흐름을 검증해 준다.

데이터셋

- CVPR 31,677 + NeurIPS 25,179 + ICML 17,059 + ICCV 12,599 + ICLR 12,265 + ECCV 11,766 + 3DV 1,638 = **112,183편**
- 1987 – 2025 (38년)
- CV+ML Paper Atlas (DBLP) 기반
- 모든 논문이 4-depth 분류됨 (Phylum/Class/Order는 100%, 미분류 5.9%)

4-depth taxonomy

L1 Phylum	16	큰 분야 (Object Detection, 3D Vision ...)
L2 Class	~120	분야 안의 갈래 (Diffusion, Pose ...)
L3 Order	~400	세부 주제 (Text-to-Image ...)
L4 Genus	가변	구체적 접근법 (Latent Diffusion ...)

왜 “생물 계통도”인가

- Linnaean taxonomy는 **진화 계통(공통 조상)**을 표현하는 트리
- CV+ML도 비슷 — 같은 “조상” 문제에서 분기
(Image Generation → GAN → Conditional GAN → StyleGAN → DDPM → Latent Diffusion → Stable Diffusion)
- 단계별 분기를 따라가면 **paradigm 전환의 시점**이 자연스럽게 드러남
- 학술적 가치 + 학생 교육자료로 적합

Methodology: semantic synonym cluster

- 단순 키워드 매칭 / TF-IDF의 한계 — 동의어를 못 잡음.
- 우리는 **rule-based + manual semantic cluster**로 명시적 분류:

예: NeRF / Neural Implicit

```
NERF = ['nerf', 'neural radiance', 'neural implicit',
        'occupancy network', 'signed distance', 'instant ngp',
        'radiance field', 'neural render', 'volumetric render', ...]
```

```
if has_any(t, NERF):
    return ('3. 3D Vision & Reconstruction',
            'Neural Implicit Representations', 'Neural Radiance Fields')
```

⇒ “Mip-NeRF”, “Block-NeRF”, “occupancy networks for 3D” 같은 표현이 **모두** 같은 leaf로 묶임. 변형(Dynamic / Human / Editing)은 Genus 레벨로 내려감.

Methodology: 우선순위 + Specific-first

- 규칙은 specific → general 순. Cross-cutting 케이스를 명시적으로 처리.
- Phylum/Class/Order는 100% 라벨링.
- Genus는 specific rule 매칭 시 (약 50%) — 나머지는 (general).
- 미분류 (Other / Unclassified) 5.9% — Position 논문, 추상적 제목, 옛 표현 등.

왜 Sentence Transformer/LLM 안 썼나

- 재현성 — rule 기반은 입력 동일하면 출력 동일.
- 투명성 — 어떤 rule이 매칭되었는지 추적 가능.
- 도메인 지식 인코딩 — CV+ML 표현 변천을 사람이 직접 큐레이션.

전체 흐름: 1987-2025 누적

- 1987-1991 cohort: **1,619편/5년** — CVPR 단독, 연 ~300편
- 2002-2006: **4,863편** (3× 성장) — NeurIPS/ICML 본격 합류
- 2012-2016: **10,846편** — AlexNet 효과 + ICLR 출범 (2013)
- 2017-2021: **31,113편** — Transformer + Self-supervised 폭발
- 2022-2025 (4년치): 이미 **50,571편** — **전체의 45%**가 이 시기

관찰

CV+ML 출판량은 38년에 걸쳐 **30× 이상** 증가.

가속은 비선형 — 2017년 이후 cohort마다 ~1.6×. **2022+는 Foundation Models 폭발.**

Phylum 분포 (전체 112,183편)

Phylum	N	%
15. Efficient & Robust ML	11,993	10.7%
12. Training Strategies	11,404	10.2%
3. 3D Vision & Reconstruction	10,970	9.8%
4. Image Recognition & Retrieval	8,090	7.2%
11. Deep Learning Architecture	7,751	6.9%
13. Optimization & Learning Theory	6,819	6.1%
7. Representation Learning	6,619	5.9%
6. Generative Models & Synthesis	6,488	5.8%
5. Video & Motion Understanding	6,215	5.5%
8. Vision-Language & Multimodal	5,691	5.1%
14. Reinforcement Learning & Decision Making	5,634	5.0%
10. Human-centric Vision	4,506	4.0%
1. Object Detection & Localization	3,859	3.4%
16. Application Domains	3,655	3.3%
2. Segmentation	2,967	2.6%
9. Low-level Vision	2,900	2.6%
Other / Unclassified	6,225	5.5%
Other / Editorial	397	0.4%

16개 Phylum이 모두 2.6–10.7% 범위 — 균형 잡힌 분포. 이전 “Training & Learning Methods” 단일 phylum이 21.3%로

Phylogeny (112,183 papers) CV-PML 38년 비대했는데, 학습 전략 / 최적화 이론 / 강화학습 셋으로 분리됨. April 28, 2026

8 cohorts × 5년 — 전체 카운트

Cohort	Papers
1987-1991	1,619
1992-1996	2,308
1997-2001	3,161
2002-2006	4,863
2007-2011	7,702
2012-2016	10,846
2017-2021	31,113
2022-2025	50,571

2017 이후 paper count가 가파르게 상승 — Transformer + ICLR 출범 + Self-supervised 폭발 + Foundation Models 효과의 합. 마지막 cohort는 4년치인데도 이미 가장 큼.

1987-1991 — 고전 CV의 시대

Headline

Geometry · MRF · low-level vision · pattern recognition. CVPR 단독 학회에서 시작 — 연 100 400편 수준.

Top-5 Classes (1,619 papers)

(첫 cohort — rank 비교 대상 없음)

1.	12. Training Strategies > Trainin...	316
2.	Other > Unclassified	178
3.	11. Deep Learning Architecture > ...	148
4.	3. 3D Vision & Reconstruction > 3...	130
5.	4. Image Recognition & Retrieval ...	124

Top-3 Phyla

12. Training Strategies (322), 4. Image Recognition & Retrieval (262), 3. 3D Vision & Reconstruction (192)

1992-1996 — 통계 학습의 등장

Headline

Boosting · Bayesian inference · early SVMs. Active vision, 비전-기반 추적이 형식화되는 시기.

Top-5 Classes (2,308 papers)

1.	12. Training Strategies > Trainin...	352
2.	3. 3D Vision & Reconstruction > 3...	243
3.	Other > Unclassified	225
4.	4. Image Recognition & Retrieval ...	166
5.	4. Image Recognition & Retrieval ...	132

급부상 (rank ↑ vs 직전 cohort)

+14	#23: 15. Efficient & Robust ML > U...	(19)
+9	#15: 5. Video & Motion Understandi...	(40)
+7	#12: 5. Video & Motion Understandi...	(44)
+7	#19: 10. Human-centric Vision > Fa...	(31)
+4	#14: 3. 3D Vision & Reconstruction...	(42)

Top-3 Phyla

12. Training Strategies (366), 3. 3D Vision & Reconstruction (345), 4. Image Recognition & Retrieval (340)

1997-2001 — Local feature 시대 개막

Headline

SIFT (1999) · SVM 보편화 · Bag-of-Visual-Words 등장. NeurIPS가 데이터셋에 본격 합류.

Top-5 Classes (3,161 papers)

1.	12. Training Strategies > Trainin...	337
2.	3. 3D Vision & Reconstruction > 3...	335
3.	Other > Unclassified	295
4.	4. Image Recognition & Retrieval ...	199
5.	4. Image Recognition & Retrieval ...	196

급부상 (rank ↑ vs 직전 cohort)

+26	#16: 13. Optimization & Learning T...	(52)
+16	#25: 4. Image Recognition & Retri...	(29)
+9	#14: 15. Efficient & Robust ML > U...	(58)
+7	#18: 5. Video & Motion Understandi...	(49)
+6	#26: 7. Representation Learning > ...	(22)

Top-3 Phyla

3. 3D Vision & Reconstruction (490), 4. Image Recognition & Retrieval (465), 12. Training Strategies (364)

2002-2006 — Object recognition + MRF 황금기

Headline

Conditional random fields · part-based models · Viola-Jones. Deformable Part Models 직전의 고전 CV 황금기.

Top-5 Classes (4,863 papers)

1.	12. Training Strategies > Trainin...	436
2.	3. 3D Vision & Reconstruction > 3...	415
3.	Other > Unclassified	398
4.	15. Efficient & Robust ML > Bayes...	339
5.	4. Image Recognition & Retrieval ...	301

급부상 (rank ↑ vs 직전 cohort)

+45	#46: 13. Optimization & Learning T...	(17)
+39	#28: 12. Training Strategies > Sem...	(33)
+20	#38: 16. Application Domains > Med...	(23)
+17	#42: 5. Video & Motion Understandi...	(18)
+16	#17: 7. Representation Learning > ...	(66)

Top-3 Phyla

4. Image Recognition & Retrieval (696), 3. 3D Vision & Reconstruction (606), 15. Efficient & Robust ML (531)

2007-2011 — Deep Learning의 씨앗

Headline

ImageNet (2009) · Deep Belief Nets · Hinton/LeCun renaissance. 아직 CNN이 CVPR 메인스트림은 아님.

Top-5 Classes (7,702 papers)

1.	3.	3D Vision & Reconstruction > 3...	633
2.	12.	Training Strategies > Trainin...	609
3.		Other > Unclassified	537
4.	15.	Efficient & Robust ML > Bayes...	449
5.	4.	Image Recognition & Retrieval ...	444

급부상 (rank ↑ vs 직전 cohort)

+37	#43:	15. Efficient & Robust ML > C...	(34)
+32	#51:	16. Application Domains > Rem...	(24)
+28	#59:	7. Representation Learning > ...	(19)
+26	#35:	6. Generative Models & Synthe...	(43)
+24	#18:	5. Video & Motion Understandi...	(105)

Top-3 Phyla

4. Image Recognition & Retrieval (1068), 3. 3D Vision & Reconstruction (986), 15. Efficient & Robust ML (793)

2012-2016 — AlexNet 혁명 → ResNet 정착

Headline

AlexNet (2012)이 ImageNet을 풀자마자 **4년 만에 모든 task 재패러다임화**. VGG (2014) · ResNet (2015) · GAN (2014) · Faster R-CNN.

Top-5 Classes (10,846 papers)

1.	12. Training Strategies > Trainin...	879
2.	Other > Unclassified	728
3.	15. Efficient & Robust ML > Bayes...	719
4.	3. 3D Vision & Reconstruction > 3...	713
5.	13. Optimization & Learning Theor...	589

급부상 (rank ↑ vs 직전 cohort)

+85	#26: 11. Deep Learning Architectur...	(85)
+54	#30: 10. Human-centric Vision > Pe...	(75)
+48	#60: 4. Image Recognition & Retri...	(30)
+33	#33: 7. Representation Learning > ...	(69)
+26	#39: 13. Optimization & Learning T...	(65)

Top-3 Phyla

4. Image Recognition & Retrieval (1283), 3. 3D Vision & Reconstruction (1146), 15. Efficient & Robust ML (1133)

2017-2021 — Transformer + Self-supervised + Diffusion

Headline

“Attention Is All You Need (2017) → ViT (2020). SimCLR/MoCo, CLIP (2021), DDPM (2020)이 동시기에 등장 — **현대 AI의 청사진 완성.**

Top-5 Classes (31,113 papers)

1.	12. Training Strategies > Trainin...	2220
2.	14. Reinforcement Learning & Deci...	1626
3.	13. Optimization & Learning Theor...	1559
4.	Other > Unclassified	1526
5.	11. Deep Learning Architecture > ...	1326

급부상 (rank ↑ vs 직전 cohort)

+95	#82: 12. Training Strategies > Kno... (76)
+78	#32: 7. Representation Learning > ... (291)
+72	#105: 11. Deep Learning Architectur... (36)
+71	#41: 11. Deep Learning Architectur... (186)
+58	#9: 6. Generative Models & Synthe... (819)

Top-3 Phyla

15. Efficient & Robust ML (3399), 12. Training Strategies (3213), 11. Deep Learning Architecture (2887)

2022-2025 — Foundation Models 시대

Headline

Stable Diffusion · LLaVA · NeRF/3DGS · Segment Anything · GPT-4V. 단 4년 만에 **50,571편** — 전체의 **45%**가 이 시기에 쏟아짐.

Top-5 Classes (50,571 papers)

1.	12. Training Strategies > Trainin...	3112
2.	8. Vision-Language & Multimodal >...	2926
3.	14. Reinforcement Learning & Deci...	2637
4.	Other > Unclassified	2338
5.	13. Optimization & Learning Theor...	2040

급부상 (rank ↑ vs 직전 cohort)

+107	#25: 8. Vision-Language & Multimod...	(540)
+97	#7: 6. Generative Models & Synthe...	(1861)
+87	#99: 7. Representation Learning > ...	(83)
+70	#59: 8. Vision-Language & Multimod...	(193)
+65	#40: 11. Deep Learning Architectur...	(348)

Top-3 Phyla

15. Efficient & Robust ML (5493), 8. Vision-Language & Multimodal (4800), 12. Training Strategies (4743)

2012-2016: AlexNet 혁명

- 2011년까지 ImageNet 챔피언은 SIFT+SVM (25% top-5 error).
- **AlexNet (2012)** — single CNN paper로 ImageNet error 절반화 (16%).
- 4년 안에 패러다임 전환:
 - VGG (2014), GoogLeNet (2014), ResNet (2015)
ResNet은 2016 CVPR 베스트 페이퍼, 인용 200k+
 - Faster R-CNN (2015), YOLO (2015), SSD (2016)
 - GAN (2014), VAE (2013) — 생성 모델 르네상스
 - FCN/SegNet/U-Net (2015) — semantic segmentation 표준
- 결론: **2012-2016이 “CNN이 모든 task를 새로 정의한” 4년.**

2017-2021: Transformer + Self-supervised + Diffusion

- “Attention Is All You Need” (2017) — NLP에서 시작했지만 CV로 빠르게 침투
- **ViT (2020)** — 16x16 patch가 CNN을 대체할 수 있음을 증명
- **Self-supervised 폭발**: SimCLR/MoCo (2020) → DINO (2021) → MAE (2022)
- **Diffusion 정착**: DDPM (2020) → Latent Diffusion (2021) → Stable Diffusion (2022)
- **CLIP (2021)** — Vision + Language의 첫 대규모 contrastive pretraining
- Detection 패러다임 전환: DETR (2020), Swin (2021)
- 이 5년이 modern AI 청사진 — 이후 시기는 이를 “규모화”.

2022-2025: Foundation Models 시대

- **Stable Diffusion (2022)** — 일반인 누구나 text-to-image.
Diffusion Models 누적 1,749편 (Image), 154편 (Video), ...
- **Segment Anything (2023)** — promptable segmentation foundation.
- **LLaVA (2023)**, GPT-4V — Visual instruction tuning + LVLM 폭발.
- **NeRF (2020) → 3D Gaussian Splatting (2023)** — 3D 표현의 분기점.
- Foundation Models가 모든 phylum의 1위 클래스를 점령:
 - Generative Models: Diffusion이 GAN을 대체
 - Vision-Language: CLIP/LVLM이 caption/VQA를 통합
 - 3D: NeRF/3DGS가 SfM/MVS를 부분 대체

사라진 분야 — 30년 시간축으로 본 “surface 검증”

- **Hand-crafted local features (SIFT/SURF/ORB)**
Pre-2012: 활발 → Post-2017: deep features (SuperPoint, LoFTR)에 흡수.
- **Bag-of-Visual-Words / Fisher Vectors**
Pre-2012: image classification 표준 → Post-2014: CNN feature가 완전 대체.
- **Deformable Part Models (DPM)**
2008-2014 detection의 SoTA → R-CNN family에 흡수.
- **고전 Conditional Random Fields (CRF)**
Pre-2015 segmentation 보조 → Post-2017 end-to-end NN에 통합.
- Taxonomy validation: 한 분야가 실제로 “죽었음”을 데이터로 확인하는 능력은 phylogenetic 관점의 핵심 가치.

학회별 색깔 (CVPR · NeurIPS · ICML · ICCV · ICLR · ECCV · 3DV)

Venue	성격	Top phyla 경향
CVPR (31,677편)	종합 CV 학회	Detection / Segmentation / 3D / Generative 골고루
NeurIPS (25,179편)	ML 메소드 중심	Training Strategies / Optimization / RL 강세
ICML (17,059편)	ML 이론·메소드	Optimization Theory / RL 강세
ICCV (12,599편)	종합 CV (격년)	CVPR과 유사
ICLR (12,265편)	표현학습 중심	Representation / DL Architecture 강세
ECCV (11,766편)	종합 CV (유럽, 격년)	CVPR과 유사
3DV (1,638편)	3D 전문	3D Vision & Reconstruction 압도

- NeurIPS/ICML/ICLR이 “ML 방법론” phyla(11-15)를 끌어올린다 — 그래서 12-15가 모두 5%+.
- CVPR/ECCV/ICCV가 vision-task phyla(1-10)를 채운다.
- **3DV는 작지만 phylum 3 (3D Vision)에 거의 100% 집중** — 분류기가 venue 정체성을 잘 잡아냄.

Phylum 16개의 “rank race” (cohort 별 1위 경쟁)

Phylum	87-91	92-96	97-01	02-06	07-11	12-16	17-21	22-25
3D Vision & Recon	1	1	1	2	3	4	6	5
Image Recognition	2	2	2	1	1	1	5	4
Object Detection	5	4	4	5	5	6	9	11
Segmentation	4	5	7	8	9	10	13	13
Training Strategies	6	6	5	4	4	3	2	2
Generative Models	—	—	—	—	8	7	8	6
Vision-Language	—	—	—	—	—	—	11	8
Foundation 3D (NeRF)	—	—	—	—	—	—	—	3

셀의 숫자는 해당 cohort 내 *Phylum rank* (1 = 최다). “-” = 의미 있는 카운트 없음.

2017+ Training Strategies가 압도적 1위는 NeurIPS/ICML 비중 증가의 결과.

교육적 활용

- **학생 진입 시**: 16 Phylum overview → 관심 Phylum의 cohort wall → 그 안의 paradigm shift를 따라가는 **독서 path**
- **문헌 조사 시**: Cohort의 “첫 논문 + dominant paper”를 자동 추천 가능 (이미 viewer에서 DOI 링크로 연결)
- **커리큘럼 설계 시**: 2022+ cohort의 “꼭 읽어야 할 N편”을 분야별로 추출 — phylogeny가 가지치기를 알려줌
- **투자/PI 의사결정**: Faded vs emerging Class의 비대칭 — 어디에 시간을 쓸 것인가

Live Tools

- Site: <https://gisbi-kim.github.io/cvml-paper-phylogeny/>
- 인터랙티브 viewer 두 가지:
 - **Radial Tree** — pie chart, 검색 + 추천 chip + DOI modal
 - **Horizontal Collapsible Tree** — 좌-우 phylogenetic 형태 + drag-pan + wheel-zoom
- URL state share: 특정 wedge 클릭 → #tab=tree&node=... URL 복사로 그 상태 공유
- KR/EN 토글: 우상단 appbar
- 데이터 refresh: REFRESH.md 참고 (DBLP에서 새 논문 끌어와서 재분류)

Limitations

- **단일 라벨**: 한 논문 = 한 카테고리. 멀티-필드 논문은 가장 specific한 곳으로 압축됨.
- **제목만 사용**: Abstract 미포함 → 제목이 모호하면 분류 정확도 저하 (Other/Unclassified **5.9%**).
- **7개 학회만**: AAAI/IJCAI/SIGGRAPH/Workshop 미포함 — 학회 우선의 분야는 과소 표현될 수 있음.
- **Citation 데이터 없음**: 본 버전은 citation 정규화 미수행. 영향력 비교는 별도 enrichment 필요.

워크 예제: Image Generation의 30년

Cohort	Top Generative Class	대표 방법론
1987-2006	거의 없음	텍스처 합성, image quilting
2007-2011	RBM / DBN	통계 모델 기반 image prior
2012-2016	GAN 등장	DCGAN (2015), pix2pix (2016)
2017-2021	GAN 정점 + Diffusion 등장	StyleGAN (2018-19), DDPM (2020), CLIP (2021)
2022-2025	Diffusion 압승	Stable Diffusion, Imagen, DALL-E 3, Sora

한 Phylum 안에서 Class 단위로 paradigm shift가 4번 — phylogeny가 그 가지치기를 추적한다.

워크 예제: 3D Vision의 30년

Cohort	State of the art
1987-1996	Stereo, Shape from X (shading/texture/contour)
1997-2006	SIFT-based MVS, structure from motion 정형화
2007-2011	PMVS, real-time SfM (Bundler)
2012-2016	학습 기반 single-view depth (Eigen 2014), MVSNet (2018)
2017-2021	NeRF (2020) — neural rendering의 시작
2022-2025	3D Gaussian Splatting + neural SDF + diffusion-based 3D

3D는 cohort마다 “representation 갱신” 패턴 — 단순 → SfM → 학습 → neural field → Gaussian.
Phylogenetic tree로 자연스럽게 표현됨.

Future Work

- **Citation enrichment**: OpenAlex/SemanticScholar API로 citation 채워서 “인기 vs 영향력” 분리
- **Cohort × Phylum heatmap**: 각 cohort의 phylum signature를 한 행으로
- **Class-level alluvial**: 인접 cohort 간 rank 흐름을 Sankey로 시각화
- **Author network 결합**: 누가 어떤 paradigm shift를 끌고 왔는가
- **AAAI/IJCAI/SIGGRAPH 통합**: 112k → 250k+ 규모로 확장

한 줄 요약

CV+ML 38년은

고전 CV (1987-2011) → AlexNet 혁명 (2012-16) → Transformer
+ SSL + Diffusion (2017-21) → Foundation Models (2022+)

4번의 분기점으로 정리된다.

Q&A

?

<https://gisbi-kim.github.io/cvml-paper-phylogeny/>