

오지형 교수

중앙대학교 첨단영상대학원

Ⅰ. 저자 소개

오지형 교수는 현재 중앙대학교 첨단영상대학원 조교수로 재직하며 Creative Vision and Multimedia Lab (CMLab)을 운영하고 있다. KAIST 전기및전자공학부에서 학사, 석사, 박사학위를 받았으며, 박사과정 중에는 미국 Meta Reality Labs에서 AI Research Scientist Intern으로 근무하며 대규모 움직임을 포함하는 비디오의 Dynamic NeRF 및 novel-view synthesis를 연구하였다. 이후 KAIST에서 박사후연구원으로 Dynamic NeRF와 Low-Level Vision 연구를 수행한 뒤, 2023년 9월 중앙대학교에 부임하였다. 현재는 영상 AI 기업 inshorts의 Chief AI Officer (CAIO)를 겸직하며 연구실에서 개발하는 Visual AI 기술과 실제 콘텐츠·미디어 환경 사이의 접점을 함께 탐구하고 있다.

CMLab이 지향하는 연구의 큰 방향은 “Generative Visual AI from Pixels to 4D Worlds”라는 문장으로 요약할 수 있다. 이미지의 픽셀 단위 품질을 복원하는 전통적인 Low-Level Vision에서 출발했지만, 연구의 범위는 자연스럽게 비디오와 움직임, 3D/4D 장면 표현, 그리고 최근의 생성형·에이전틱 AI로 확장되어 왔다. 서로 다른 연구 분야를 단순히 병렬적으로 다루기보다는, **AI가 시각 정보를 더 잘 보고, 복원하고, 시간적 움직임과 공간적 구조를 이해하며, 궁극적으로는 사용자의 의도에 따라 생성하고 편집할 수 있도록 하는 하나의 Visual Intelligence 문제**로 바라보고 있다.

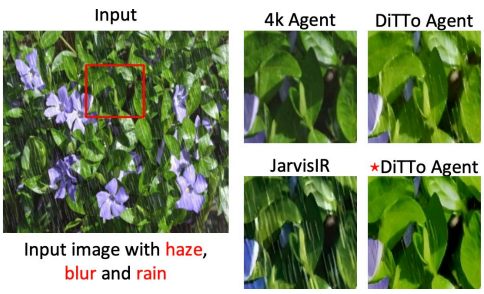


Ⅱ. 연구 분야

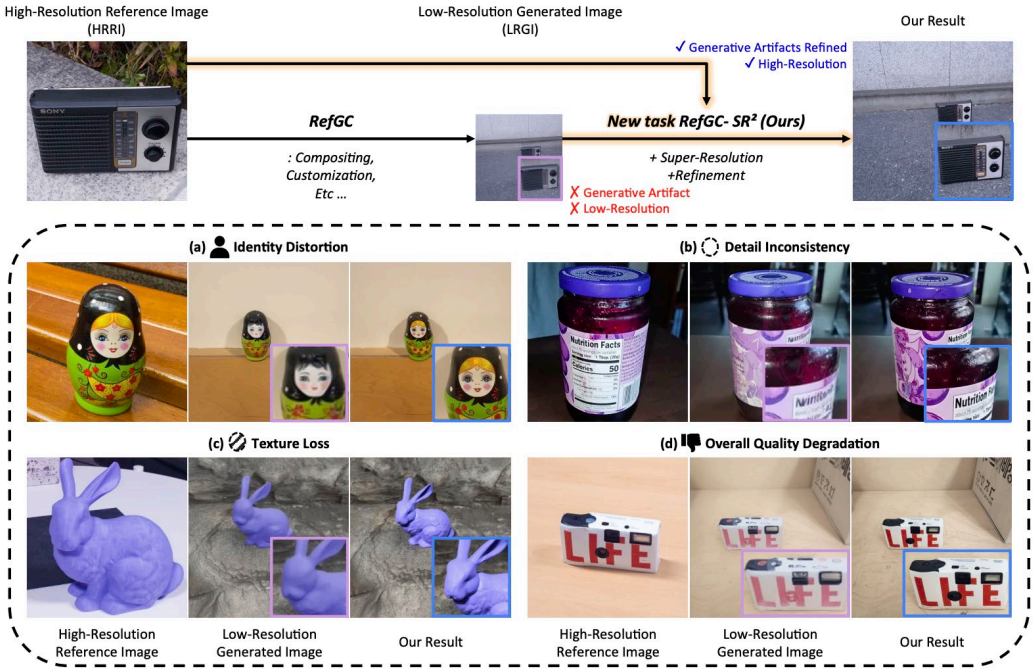
1. Pixel-Level Intelligence

이미지 복원과 화질 개선은 CMLab의 가장 오래된 연구 기반 중 하나이다. Super-Resolution, Deblurring, Demoiréing과 같이 손상되거나 부족한 시각 정보를 복원하는 문제는 오랫동안 연구되어 왔지만, 최근 생성형 AI와 foundation model의 등장으로 연구 질문 자체가 빠르게 변화하고 있다. 기존에는 하나의 degradation에 특화된 모델을 설계하고 PSNR이나 SSIM과 같은 정량 지표를 개선하는 데 초점을 두었다면, 최근에는 **실세계에서 발생하는 복합적인 열화를 어떻게 진단할 것인지, 여러 복원 과정을 어떤 순서로 수행할 것인지, 생성형 AI가 만들어낸 콘텐츠를 어떻게 신뢰성 있게 다시 정제할 것인지**가 중요해지고 있다.

CMLab에서는 이러한 흐름을 바탕으로 real-world super-resolution, joint video restoration, all-in-one image restoration뿐 아니라 **AIGC refinement와 agentic restoration**까지 연구 영역을 확장하고 있다. 예를 들어 FRAMER는 diffusion prior와 frequency-aligned self-distillation을 활용하여 실제 이미지의 super-resolution 성능을 개선하였으며 CVPR 2026에 발표되었다. 한편 DiTTo는 여러 열화가 동시에 존재하는 입력에서 복원의 순서까지 고려하는 All-in-One Image Restoration Agent를 연구한다. 또한 Adobe Research와의 공동연구에서는 reference information을 활용해 저해상도 또는 생성형 AI 콘텐츠를 정교하게 복원하는 RefGC-SR²와 cross-domain object compositing 등을 연구하며,



<DiTTo: Scalable Order-aware All-in-One Image Restoration Agent>



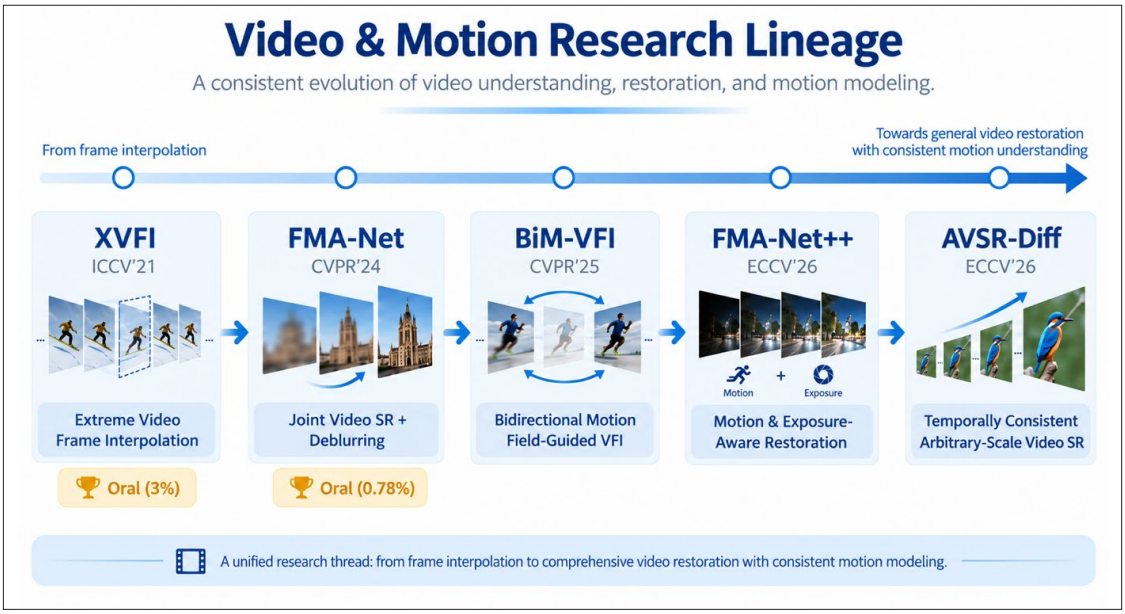
<RefGC-SR²: Reference-guided Super-Resolution and Refinement of AI Generated Content>

전통적인 Low-Level Vision을 **생성형 시각 콘텐츠의 품질 제어 문제**까지 확장하고 있다.

2. Video & Motion Intelligence

정지영상과 달리 비디오는 시간에 따른 움직임이라는 추가적인 정보를 갖는다. 이는 복원과 생성에 유용한 단서를 제공하는 동시에, temporal inconsistency나 motion artifact와 같은 새로운 문제를 만들어낸다.

오지형 교수는 대학원 과정부터 Video Frame Interpolation과 Video Restoration을 지속적으로 연구해 왔다. Extreme motion을 다룬 **XVFI**는 ICCV 2021 Oral (**3%**)로 선정되었고, Video Super-Resolution과 Deblurring을 하나의 frame-



work에서 동시에 해결한 **FMA-Net**은 CVPR 2024 Oral (**0.78%**)로 선정되었다. 이후 BiM-VFI, FMA-Net++, AVSR-Diff 등으로 연구를 확장하며 non-uniform motion, exposure variation, arbitrary-scale video restoration과 temporal consistency를 지속적으로 다루고 있다. XVFI와 FMA-Net은 각각 ICCV와 CVPR에서 Oral 논문으로 선정된 연구들이다.

최근에는 그동안 축적된 Video Frame Interpolation 연구를 체계적으로 정리한 **AceVFI**가 IEEE TCSVT에 게재되었으며, 개별 frame의 품질을 향상시키는 것을 넘어 **비디오 전체에서 움직임을 어떻게 표현하고 유지할 것인가**라는 방향으로 연구가 확장되고 있다.

이러한 video 및 motion 연구는 이후 4D scene modeling이나 video generation에서도 중요한 기반이 된다. CMLab에서는 장기적으로 2D 비디오의 motion reasoning과 4D world representation 사이의 경계를 점차 연결하고자 한다.

3. 3D/4D World Representation

최근 Computer Vision은 2D 이미지에서 객체를 인식하는 수준을 넘어, 실제 세계의 3차원 구조와 시간적 변화를 표현하는 방향으로 빠르게 확장되고 있다. Neural Radiance Fields와 Gaussian Splatting의 등장으로 복잡한 장면을 높은 품질로 표현하는 것이 가능해졌지만, **동적인 장면, 긴 시간 범위의 움직임, 실제 촬영 환경의 blur, 효율적인 표현과 렌더링**은 여전히 중요한 연구 문제이다.

CMLab에서는 Dynamic NeRF, 3D Gaussian Splatting, Diffusion, VGGT 계열 모델들 기반으로 이러한 문제들을 연구하고 있다. MoDec-GS와 SplineGS에서는 동적 장면의 motion representation을 다루었으며, **MoRel**에서는 긴 시간 범위에 걸쳐 변화하는 장면에서 flicker를 억제하면서 안정적으로 4D motion을 모델링하는 문제를 연구하였다. MoRel은 CVPR 2026 Highlight (**3.6%**) 논문으로 선정되었다. 이와 함께 feed-forward 3D Gaussian Splatting의 efficiency를 조절하는 **EcoSplat** 역시 CVPR 2026 Highlight (**3.6%**)로 선정되었다.



연구실의 프로젝트 역시 이러한 연구 방향과 연결되어 있다. 현재 CMLab은 VLM agent를 활용한 복합 열화 진단 기반 4D 동적 장면 복원, dynamic spatial prior를 활용한 4DGS 생성, 대규모 병렬학습을 통한 feed-forward 3D scene generation 등 **3D/4D world modeling과 관련된 연구책임 프로젝트**를 진행하고 있다.

궁극적으로 관심을 두고 있는 것은 단순히 사진 몇 장으로 새로운 시점을 렌더링하는 것을 넘어, **AI가 실제 세계의 공간적 구조와 시간적 변화를 표현하고, 이를 생성·편집·추론에 활용할 수 있는 Physical World Model**로의 확장이다.

4. Generative & Agentic Visual AI

Diffusion model과 multimodal foundation model의 발전으로 생성형 AI는 이미지를 새롭게 생성하는 단계를 넘어, 사용자의 의도를 이해하고 기존 콘텐츠를 정교하게 편집하는 방향으로 발전하고 있다.

CMLab에서는 image morphing, personalization, object compositing, text-image manipulation, multi-object video editing과 같은 **controllable visual generation** 문제를 연구하고 있다. OmniText는 별도의 추가학습 없이 다양한 text-image manipulation task를 수행하는 generalist를 제안하여 ICLR 2026에 발표되었으며, PRIMEdit은 multi-object video editing을 위한 새로운 benchmark와 editing framework를 제시하여 IEEE TPAMI에 게재되었다. 최근에는

diffusion transformer 기반의 multi-object motion transfer를 다루는 MotionGrounder가 NeurIPS 2026에 발표되었다.

동시에 연구 관심은 단순한 생성 모델에서 **Agentic Visual AI**로 이동하고 있다. 예를 들어 image restoration에서는 agent가 입력의 degradation을 판단하고 적절한 처리 순서를 선택할 수 있으며, 향후에는 editing, restoration, generation, 3D reconstruction 등의 여러 visual tool을 목적에 따라 조합하는 방식으로 확장할 수 있다.

이 과정에서 CMLab은 Adobe Research, Princeton University, KAIST, Kyung Hee University를 비롯한 국내외 연구기관 및 산업체 연구자들과 다양한 형태의 공동연구를 진행하고 있다. 공개된 최근 연구에는 Adobe Research와의 compositing 및 AIGC refinement, Princeton University와의 morphing 및 3D low-level vision, Kyung Hee University와의 3D vision 연구 등이 포함되어 있다.

이러한 공동연구는 서로 다른 문제와 관점을 연결해 **한 연구실 내부에서 해결하기 어려운 새로운 문제를 정의하는 과정**이라고 생각한다.

Ⅲ. 연구에서 실제 시스템으로의 확장: Real-World AI Systems & Industry Translation

오지형 교수가 연구실 운영에서 중요하게 생각하는 또 하나의 축은 **학술적인 novelty와 실제 활용 가능성을 분리하지 않는 것**이다.

2024년부터 AI 기반 영상 기술 기업인 inshorts에서 CAIO를 겸직하며 Low-Level Vision 기술의 연구 방향 설정과 실제 제품 적용을 함께 고민하고 있다. 실제 산업 현장에서는 benchmark에서의 작은 성능 차이보다, 예상하지 못한 입력, 계산 비용, latency, temporal stability, 콘텐츠 제작자의 요구 등 논문에서 쉽게 드러나지 않는 문제가 훨씬 중요하게 작용한다. 이러한 현실적인 제약은 반대로 새로운 연구 문제를 발견하는 좋은 출발점이 되기도 한다.

CMLab에서는 정부 연구과제뿐 아니라 ETRI, DEXTER STUDIOS 등과의 연구용역과 GPU 컴퓨팅 지원사업을 통해 3D/4D generation, style transfer, agentic restoration, feed-forward 3D generation과 같은 연구를 수행하고 있다. 동시에 한국연구재단의 중견연구 과제를 통해 **User-Interactive Next-Generation AI for Multi-Modal Immersive Content Technology**를 연구하고 있다.

이러한 구조가 학계와 산업계의 협업에서 특히 중요한 장점이라고 생각한다. 기업이 보유한 실제 데이터와 서비스 환경에서 발생하는 어려운 문제를 연구자가 새로운 benchmark나 algorithmic problem으로 정식화하고, 다시 연구 결과를 실제 시스템으로 검증하는 **양방향 연구 구조**를 만들 수 있기 때문이다.

따라서 CMLab은 단순한 기술 자문뿐 아니라, **새로운 문제 정의, 데이터 및 benchmark 설계, 핵심 AI 모델 개발, 실제 시스템 검증**까지 이어지는 공동연구를 지향한다. 방송·미디어, 콘텐츠 제작, 영상 품질 개선, 생성형 AI, 3D/4D 콘텐츠, multimodal/agentic AI와 관련된 다양한 산학협력에 열려 있다.

Ⅳ. 최근 연구 성과

CMLab은 2023년 가을 출범한 비교적 젊은 연구실이다. 그러나 연구실의 성장 과정에서 중요한 것은 단순히 논문

의 수를 늘리는 것이 아니라, 앞서 소개한 네 개의 연구축이 점차 서로 연결되는 연구 포트폴리오를 만드는 것이었다.

2026년 10월 기준으로, 관련 연구는 **28편의 peer-reviewed international publication**으로 이어졌으며, 이 가운데 **20편이 주요 top-tier AI/CV 국제학회 논문이고 8편이 국제저널 논문**이다. 이외에도 10편의 연구가 preprint 또는 심사 중이다. CVPR에 8편, ICCV에 3편, ECCV에 3편을 발표했으며, NeurIPS, ICLR, AAAI 및 IEEE TPAMI, IEEE TCSVT 등으로 연구 영역을 확대하고 있다.

특히 최근에는 CVPR 2026에서 MoRel과 EcoSplat이 Highlight로 선정되었고, 2026년에는 FMA-Net++과 AVSR-Diff의 ECCV 발표, MotionGrounder의 NeurIPS 발표, PRIMEdit의 IEEE TPAMI 게재, AceVFI의 IEEE TCSVT 게재 등으로 이어졌다.

이러한 성과에서 주목해야할 점은 **학생들이 각자의 독립적인 연구 주제를 만들기 시작했다는 점**이다. 연구실 내부에서는 서로 다른 task를 수행하는 학생들이 정기적으로 연구 방향을 공유하고, motion, restoration, 3D/4D, generative AI처럼 서로 연결될 수 있는 주제를 발견하면 공동연구로 발전시키고 있다.

또한 국제학회 발표, 워크숍 운영, 국내외 공동연구 등을 통해 학생들이 연구실 내부에 머무르지 않고 보다 넓은 연구 커뮤니티와 연결될 수 있도록 노력하고 있다. 최근에는 ICIP 2026 Satellite Workshop인 Many Lenses, One World (MLOW)의 Lead Organizer를 맡는 등 연구 결과를 발표하는 것을 넘어 새로운 연구 주제와 커뮤니티를 만들어가는 활동에도 참여하고 있다.

저 자 소 개



오 지 형

- 2017년 : KAIST 전기및전자공학부 학사
- 2019년 : KAIST 전기및전자공학부 석사
- 2023년 : KAIST 전기및전자공학부 박사
- 2022년 : Meta Reality Labs, AI Research Scientist Intern
- 2023년 : KAIST 정보전자연구소 박사후연구원
- 2023년 9월 ~ 현재 : 중앙대학교 첨단영상대학원 조교수 / CMLab Director
- 2024년 11월 ~ 현재 : inshorts Chief AI Officer (CAIO), 겸직
- 주관심분야 : Pixel-Level Intelligence, Video & Motion Intelligence, 3D/4D World Modeling, Generative & Agentic Visual AI, Real-World AI Systems & Industry Translation