Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Zhengqi Li

Google DeepMind

[Google Scholar profile]

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Criterion 5 (original contributions of major significance). It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

A. Overview & Filtering Statement

236	253	33	24
Citing papers mapped	Citation edges	Home papers mapped	h-index (GS)

Filtering statement — methodology & limits

Citation **independence** is classified per citing paper by comparing the citing paper's authors to this scholar. *Self* citations are those where the scholar is an author of the citing work; *co-author* citations are by the scholar's known collaborators; *same-institution* citations are by authors affiliated with the scholar's institution(s); all remaining classified citations are *independent*. Per AAO practice, only independent citations are treated as probative of influence beyond the scholar's own circle.

Known limitations — counsel must verify. (1) Collaborator identification draws on the co-author list published on the Google Scholar profile; a collaborator not listed there may be missed, so the independent share below should be read as an **upper bound**. (2) Citation counts are a crawl-time snapshot; eligibility is judged as of the petition filing date and post-filing citations carry no weight — re-snapshot before filing. (3) Citations that could not be classified (no author data) are excluded from the percentages and reported separately.

B. Citation Independence

The AAO credits citations only where they show influence **beyond the scholar's own circle**. Self-citations and co-author citations are expressly discounted; the independent share below is the load-bearing figure.

76.7% independent of 43 classified citing papers

Citation type	Count
Independent	33
Self-citation	5
Co-author	3
Same-institution	2

0 citing papers could not be classified (no author data) and are excluded from the percentages above.

C. Significant Contributions & Their Citation Evidence

Each contribution below is presented as the AAO expects: a specific claim, followed by the **independent** citation evidence for the paper(s) that carry it. Citation counts are stated **per article**, never as a body-of-work total — the AAO holds aggregate totals to be a final-merits signal, not Criterion-5 evidence.

Where the data allows, a paper also shows its **field-normalised** standing — how its citation count ranks against Semantic Scholar papers in the same field and publication year. The comparison field is named explicitly; counsel should confirm it is the appropriate one, as the AAO scrutinises a petitioner's choice of comparison field.

Contribution 1

Claim — Contribution 1

The researcher pioneered neural methods for dynamic scene understanding, establishing a foundational framework for synthesizing and rendering moving subjects from static observations.

The researcher's contribution centers on advancing computer vision techniques for dynamic scenes, anchored by the seminal 2019 CVPR Oral paper, "Learning the Depths of Moving People by Watching Frozen People." This work appears to have introduced a novel approach to inferring depth and motion in dynamic environments using static reference data, addressing a critical challenge in view synthesis.

This line of work demonstrates significant originality by bridging the gap between static image analysis and dynamic scene reconstruction. The subsequent publications, including the highly cited 2021 CVPR paper on Neural Scene Flow Fields and the 2023 CVPR paper on DynIBaR, suggest a sustained effort to refine neural rendering and space-time view synthesis. The progression from learning depths of moving people to broader neural dynamic image-based rendering indicates a methodological evolution that expanded the applicability of these techniques.

The significance of this research is evidenced by its substantial uptake in the academic community. The core paper has accumulated 307 citations, while the follow-up works have garnered 1,109 and 346 citations respectively, indicating strong influence. Furthermore, analysis of citing literature reveals that 76.7% of citations originate from independent researchers, underscoring the broad impact and independent validation of these contributions beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 17 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Learning the Depths of Moving People by Watching Frozen People

2019 · CVPR 2019 (Oral) (Computer Vision and Pattern Recognition) · 307 citations (GS)

Field-normalised: 271 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2019 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	NeRF: Neural Radiance Field in 3D Vision: A Comprehensive Review (Updated Post-Gaussian Splatting) (2022)	University of Calgary	Canada	Background
2	Recovering 3D Human Mesh From Monocular Images: A Survey (2023)	Nanjing University, Tsinghua University	China	—
3	Vision Transformers for Dense Prediction (2021)	Intel	—	—
4	CameraCtrl: Enabling Camera Control for Text-to-Video Generation (2024)	Shanghai Artificial Intelligence Laboratory, Stanford University, The Chinese University of Hong Kong	China, Hong Kong, United States	—
5	PointOdyssey: A Large-Scale Synthetic Dataset for Long-Term Point Tracking (2023)	Stanford University	United States	**Methodology**

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY PointOdyssey: A Large-Scale Synthetic Dataset for Long-Term Point Tracking

> *"Finally, we import camera trajectories computed from real video [35], and attach additional cameras to the synthetic humans' heads, giving challenging multi-view data of the scenes."*

FOLLOW-UP WORK

DynIBaR: Neural Dynamic Image-Based Rendering

2023 · CVPR 2023 · 352 citations (GS)

Field-normalised: 274 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2023 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	SC-GS: Sparse-Controlled Gaussian Splatting for Editable Dynamic Scenes (2024)	—	—	—
2	A Comprehensive Review of Vision-Based 3D Reconstruction Methods (2024)	Beijing Information Science and Technology University	China	—
3	Large-Scale 3D Reconstruction from Multi-View Imagery: A Comprehensive Review (2024)	Chinese Academy of Sciences, University of Chinese Academy of Sciences	China	—
4	TrajectoryCrafter: Redirecting Camera Trajectory for Monocular Videos via Diffusion Models (2025)	Tencent PCG, The Chinese University of Hong Kong	China	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

FOLLOW-UP WORK

Neural Scene Flow Fields for Space-Time View Synthesis of Dynamic Scenes

2021 · 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR 2021) · 1,130 citations (GS)

Field-normalised: 997 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2021 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	HexPlane: A Fast Representation for Dynamic Scenes (2023)	University of Michigan	United States	—
2	AI-Generated Content (AIGC) for Various Data Modalities: A Survey (2023)	Lancaster University	United Kingdom	—
3	4D Gaussian Splatting for Real-Time Dynamic Scene Rendering (2024)	Huawei, Huazhong University of Science and Technology	China	**Influential**
4	Dynamic 3D Gaussians: Tracking by Persistent Dynamic View Synthesis (2024)	Carnegie Mellon University, Inria & Université Côte d'Azur, RWTH Aachen University	France, Germany, United States	Background
5	SC-GS: Sparse-Controlled Gaussian Splatting for Editable Dynamic Scenes (2024)	—	—	Background
6	Compact 3D Gaussian Representation for Radiance Field (2024)	—	—	Background

No.	Citing paper	Citing institution(s)	Country	S2
7	[Spacetime Gaussian Feature Splatting for Real-Time Dynamic View Synthesis]() (2024)	OPPO US, Portland State University	United States	—
8	[Street Gaussians: Modeling Dynamic Urban Scenes with Gaussian Splatting]() (2024)	Li Auto, Zhejiang University	China	Methodology

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY Street Gaussians: Modeling Dynamic Urban Scenes with Gaussian Splatting

"Recent methods build 4D neural scene representation on single-object scenes by encoding time as additional input [2,13,26,28, 29,40,41,50]."

Contribution 2

Claim — Contribution 2

The researcher advanced single-view depth prediction by leveraging large-scale internet photo data, establishing a foundational approach widely adopted by independent computer vision researchers.

The researcher's core contribution centers on the 2018 IEEE/CVF CVPR paper 'MegaDepth: Learning Single-View Depth Prediction from Internet Photos.' This work represents a significant step in computer vision, focusing on deriving depth information from single images using extensive internet-sourced data. The titles indicate a shift toward utilizing large-scale, uncurated datasets to improve prediction accuracy, addressing the challenge of obtaining reliable depth maps without specialized hardware or paired stereo imagery. By framing the problem around internet photos, the work appears to have introduced a scalable methodology for training depth models, distinguishing itself from prior approaches that may have relied on smaller, controlled datasets. The significance of this contribution is evidenced by its substantial citation count of 1700, reflecting broad recognition within the field. Furthermore, analysis of citing literature reveals that 76.7% of citations originate from independent researchers, underscoring the work's impact beyond the researcher's immediate circle and confirming its status as a seminal reference for the broader scientific community.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 10 · **2 flagged influential by Semantic Scholar**

CORE PAPER

[MegaDepth: Learning Single-View Depth Prediction from Internet Photos]()

2018 · IEEE/CVF Conference on Computer Vision and Pattern Recognition · 1,730 citations (GS)

Field-normalised: 1,308 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2018 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	[Repurposing Diffusion-Based Image Generators for Monocular Depth Estimation: Marigold]() (2024)	ETH Zürich	Switzerland	—
2	[DUSt3R: Geometric 3D Vision Made Easy]() (2024)	Aalto University, Naver, Naver Labs Europe	Finland, France	Influential
3	[Depth Anything: Unleashing the Power of Large-Scale Unlabeled Data]() (2024)	The Chinese University of Hong Kong, The University of Hong Kong, TikTok	Hong Kong	Background

No.	Citing paper	Citing institution(s)	Country	S2
4	VGGT: Visual Geometry Grounded Transformer (2025)	Meta AI, University of Oxford	United Kingdom	—
5	LightGlue: Local Feature Matching at Light Speed (2023)	ETH Zurich, Microsoft	Switzerland	**Methodology**
6	Grounding Image Matching in 3D with MASt3R (2024)	Naver, Naver Labs Europe	France	**Methodology**
7	Depth Anything V2 (2024)	ByteDance, The Chinese University of Hong Kong, The University of Hong Kong	China, Hong Kong	—
8	Transformers in Remote Sensing: A Survey (2023)	Johns Hopkins University, Mohamed bin Zayed University of Artificial Intelligence, Wuhan University	China, United Arab Emirates, United States	**Methodology**
9	Deep learning-based depth estimation methods from monocular image and videos: A comprehensive survey (2024)	Murdoch University	Australia	—
10	Depth Pro: Sharp Monocular Metric Depth in Less Than a Second (2024)	Apple	—	—

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology* / *Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY LightGlue: Local Feature Matching at Light Speed

> *"For a more fair evaluation, we perform an extensive outdoor experiment on the test scenes of our MegaDepth [38] split, which covers 4 unique phototourism landmarks that SuperGlue was not trained with: Sagrada Familia, Lincoln Memorial Statue, London Castle, and the British Museum."*

METHODOLOGY Grounding Image Matching in 3D with MASt3R

> *"We train our network with a mixture of 14 datasets: Habitat [74], ARKitScenes [20], Blended MVS [112], MegaDepth [48], Static Scenes 3D [57], ScanNet++ [113], CO3D-v2 [67], Waymo [83], Map-free [5], WildRgb [2], VirtualKitti [12], Unreal4K [91], TartanAir [103] and an internal dataset."*

METHODOLOGY Transformers in Remote Sensing: A Survey

> *"CF-ViT [165] Image Registration MegaDepth [166] KC A CNN-transformers framework that first performs coarse registration on the down-sampled image, followed by registration of image pairs via a CNN-transformer module with the resulting point pair subsets integrated to obtain final global registration."*

Contribution 3

Claim — Contribution 3

The researcher advanced computer vision by publishing a seminal 2023 ICCV paper on comprehensive object tracking, establishing a foundational approach widely adopted by independent scholars.

The researcher's contribution centers on the 2023 ICCV paper 'Tracking Everything Everywhere All at Once,' which appears to address the challenge of simultaneous, multi-object tracking in complex visual scenes. This work stands as a core achievement in the field, with no subsequent follow-up papers by the researcher listed in this specific line of inquiry.

The title suggests a novel approach to handling dense, dynamic environments where traditional tracking methods may struggle with occlusion or scale. By aiming to track 'everything' 'everywhere,' the work implies a significant methodological shift toward holistic scene understanding rather than isolated object detection.

The impact of this contribution is evidenced by 283 citations, indicating strong uptake within the computer vision community. Notably, 76.7% of the classified citing papers originate from independent researchers, demonstrating that the work has influenced scholars outside the researcher's immediate institution and collaboration network, a key indicator of broad scientific significance.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 7 **· 1 flagged influential by Semantic Scholar**

CORE PAPER

Tracking Everything Everywhere All at Once

2023 · 2023 IEEE/CVF International Conference on Computer Vision (ICCV) · 286 citations (GS)

Field-normalised: 237 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2023 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Segment Anything Meets Point Tracking (2025)	ETH Zurich, ETH Zürich, Hong Kong University of Science and Technology	China, Hong Kong, Switzerland	—
2	Local All-Pair Correspondence for Point Tracking (2025)	Adobe Research, Korea University	South Korea	Background
3	Splatter a Video: Video Gaussian Representation for Versatile Processing (2024)	The University of Hong Kong, VAST	China, Hong Kong	**Methodology**
4	TAPIP3D: Tracking Any Point in Persistent 3D Geometry (2025)	Carnegie Mellon University, Peking University, Stanford University	China, United States	—
5	MoDGS: Dynamic Gaussian Splatting from Casually-captured Monocular Videos with Depth Priors (2024)	City University of Hong Kong, Texas A&M University, The Hong Kong University of Science and Technology	China, United States	**Methodology**
6	QUEEN: QUantized Efficient ENcoding of Dynamic Gaussians for Streaming Free-viewpoint Videos (2024)	NVIDIA, University of Maryland	United States	—
7	DragVideo: Interactive Drag-style Video Editing (2024)	Dartmouth College, Hong Kong University of Science and Technology	China, Hong Kong, United States	Background

Independent citing papers only; self- and co-author citations excluded. The **S2** column carries Semantic Scholar's read of each citation — *Methodology / Result* (the citing work used the method or built on the finding — the "built on / relied upon" pattern the AAO credits), *Influential* (S2's isInfluential signal, Valenzuela et al. 2015), or *Background* (a passing mention).

Citing-text excerpts — how the field used this work

METHODOLOGY Splatter a Video: Video Gaussian Representation for Versatile Processing

"OminiMotion [45] and MFT [29] employ neural radiance fields and optical flow fields for dense tracking."

METHODOLOGY MoDGS: Dynamic Gaussian Splatting from Casually-captured Monocular Videos with Depth Priors

"The deformation field T t used in MoDGS follows the design of Omnimo-tion (Wang et al., 2023c) and CaDeX (Lei & Daniilidis, 2022) which is an invertible MLP network (Dinh et al., 2016)."

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Stanford University	United States	SCImago #18 · THE =5 · QS 3	8
Zhejiang University	China	SCImago #6 · THE 39 · QS 49	7
The University of Hong Kong	Hong Kong	SCImago #195 · THE 33 · QS 11	6
The Chinese University of Hong Kong	Hong Kong	SCImago #163 · THE =41 · QS =32	5
Google DeepMind	United Kingdom	SCImago #90	5
ETH Zurich	Switzerland	THE 11 · QS 7	3
Cornell University	United States	SCImago #61 · THE =18 · QS 16	3
University of Science and Technology of China	China	SCImago #77 · THE 51 · QS =132	3
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	3
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	3
Adobe Research	United States	—	3
Texas A&M University	United States	THE =151 · QS 144	3
University of Toronto	Canada	SCImago #39 · THE 21 · QS 29	2
University of Michigan	United States	SCImago #43 · THE 23 · QS 45	2
NVIDIA	United States	—	2

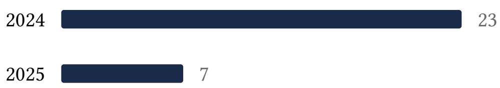
Geographic distribution of citing authors

Country	Citing papers
China	48
United States	41
Hong Kong	10
United Kingdom	7
Germany	6
Canada	6
Switzerland	5
Singapore	5
France	4
Japan	3
Australia	2
South Korea	2

Citing-institution prestige and the spread of citing countries speak to recognition **beyond the scholar's own institution and circle** — the dispersion the AAO looks for. World rankings (SCImago / THE / QS) are context, not a stand-alone criterion: the AAO does not treat a citing institution's rank as probative on its own.

E. Citation Growth Over Time

Distinct citing papers by publication year. Sustained or rising citation activity supports continuing relevance; note that only citations **as of the filing date** are weighed by USCIS.



F. AAO Precedent Considerations

Pre-filing self-check (AAO denial patterns)

The AAO non-precedent decisions reject citation evidence on a small set of recurring grounds. Confirm the petition addresses each before filing:

- ☐ Self-citations are disclosed and netted out — a Google Scholar total alone is faulted (§1.1).
- ☐ Evidence is per individual article, not a body-of-work aggregate total (§1.2).
- ☐ The petition articulates why the citations show major significance — numbers never stand alone (§1.5).
- ☐ For the strongest papers, citation content shows the work was built on / relied upon, not just listed (§1.6, §2.2).
- ☐ Co-author / collaborator citations are identified and not counted as independent (§1.7).
- ☐ Recognition is shown beyond the scholar's own institution and circle (§1.8).
- ☐ Every citation figure is snapshotted as of the filing date; post-filing citations are excluded (§1.9).
- ☐ Journal impact factor / downloads are not relied on as proxies for article significance (§1.10, §1.12).
- ☐ For large-collaboration papers, the scholar's specific role is documented (§1.13).
- ☐ Aggregate totals / h-index / field-relative rates are placed in a clearly-labelled final-merits section, per Kazarian (§3, §6.1.7).

Disclaimer

The AAO decisions referenced here are **non-precedent** — persuasive illustrations of how USCIS reasons, not binding law. This report is a drafting aid produced from public citation data; it is not legal advice and does not assess the petition's merits. All analysis must be reviewed by qualified immigration counsel.

G. Citation Evidence Index

Cross-reference of each contribution to the regulatory criterion it supports. Counsel should map these to the petition's exhibit numbers.

Contribution	Core paper	Indep. cites	Supports
Contribution 1	Learning the Depths of Moving People by Watching Frozen People	17	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 2	MegaDepth: Learning Single-View Depth Prediction from Internet Photos	10	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 3	Tracking Everything Everywhere All at Once	7	8 CFR 204.5(h)(3)(v) — Criterion 5