Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

8 CFR § 204.5(i)(3) · Authorship + Original Contributions

Tinghui Zhou

Roblox, Foundation AI

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. 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

4,785	4,887	30	19
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.

96.7% independent of 4,040 classified citing papers

Citation type	Count
Independent	3,908
Self-citation	3
Co-author	129
Same-institution	0

745 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 unsupervised methods for learning depth and ego-motion from video, establishing a foundational framework that enabled subsequent advances in view synthesis and multiplane image processing.

The researcher's core contribution rests on the 2017 paper 'Unsupervised learning of depth and ego-motion from video,' which appears to have introduced a novel approach to deriving spatial understanding from monocular video without labeled data. This work serves as the technical foundation for the researcher's subsequent publications, including the 2018 paper 'Stereo magnification: Learning view synthesis using multiplane images.'

This line of work appears to address the challenge of reconstructing three-dimensional scene geometry and synthesizing novel views from limited visual inputs. By moving from basic depth estimation to more complex view synthesis techniques involving multiplane images, the researcher demonstrates a logical progression in solving problems related to spatial perception and rendering in computer vision.

The significance of this contribution is evidenced by the high citation count of the core paper, which has been cited over 3,700 times. Furthermore, analysis indicates that 99.2% of these citations originate from independent researchers, suggesting that the work has had a broad and autonomous impact on the field, rather than relying on self-citation or institutional clustering.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1,695 · **199 flagged influential by Semantic Scholar**

CORE PAPER

Unsupervised learning of depth and ego-motion from video

2017 · Proceedings of the IEEE conference on computer vision and pattern ..., 2017 · 3,793 citations (GS)

Field-normalised: 2,868 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2017 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A survey of state-of-the-art on visual SLAM	University of Limerick	Ireland	—
2	Self-supervised visual feature learning with deep neural networks: A survey	City University of New York, The Graduate Center, The City University of New York	United States	Methodology
3	A comprehensive survey of loss functions and metrics in deep learning	Centro de Investigaciones en Óptica A.C., Instituto Politecnico Nacional, Universidad Autónoma de Querétaro	Mexico	—
4	Surroundocc: Multi-camera 3d occupancy prediction for autonomous driving	Tianjin University, Tsinghua University	China	Methodology
5	Vggt: Visual geometry grounded transformer	Meta AI, University of Oxford	United Kingdom, United States	—
6	Review of deep learning algorithms and architectures	University of Bridgeport	United States	Background
7	Learning to drive in a day	Wayve	United Kingdom	Background
8	Scaling and benchmarking self-supervised visual representation learning	Facebook AI Research	United States	Background
9	Deep learning for 3d reconstruction, augmentation, and registration: A review paper	University of Tartu, Yıldız Technical University	Estonia, Turkey	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
10	St-gan: Spatial transformer generative adversarial networks for image compositing	Adobe Research, Carnegie Mellon University	United States	Background
11	Memory-based deep reinforcement learning for obstacle avoidance in UAV with limited environment knowledge	Indian Institute of Science, Microsoft	India, United States	—
12	Dust3r: Geometric 3d vision made easy	Aalto University, Naver Labs Europe	Finland, France	—
13	Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer	ETH Zürich, Intel	Switzerland	Background
14	Vggsfm: Visual geometry grounded deep structure from motion	Meta AI, University of Oxford	United Kingdom, United States	—
15	Lite-mono: A lightweight cnn and transformer architecture for self-supervised monocular depth estimation	University of Twente	Netherlands	**Methodology**
16	Pointconv: Deep convolutional networks on 3d point clouds	Oregon State University, Tencent PCG	United States	—
17	Unifying flow, stereo and depth estimation	Beihang University, ETH Zurich, Monash University	Australia, China, Germany	**Methodology**
18	Nope-nerf: Optimising neural radiance field with no pose prior	Mohammed Bin Zayed University of Artificial Intelligence, University of Oxford	United Arab Emirates, United Kingdom	Background
19	Barf: Bundle-adjusting neural radiance fields	Carnegie Mellon University, Carnegie Mellon University, The University of Adelaide, Cornell University	United States	Background
20	Monovit: Self-supervised monocular depth estimation with a vision transformer	East China University of Science and Technology, Phi-Gent Robotics, University of Bologna	China, Italy	**Methodology**
21	Hrda: Context-aware high-resolution domain-adaptive semantic segmentation	ETH Zurich, ETH Zurich; KU Leuven	Switzerland, Switzerland; Belgium	—
22	Selfocc: Self-supervised vision-based 3d occupancy prediction	Tsinghua University	China	**Methodology**
23	Geonet: Unsupervised learning of dense depth, optical flow and camera pose	SenseTime Research	China	**Methodology**
24	Robust dynamic radiance fields	KAIST, Meta, National Taiwan University	South Korea, Taiwan, United States	**Methodology**
25	Multimodal fusion and vision-language models: A survey for robot vision	Beijing University of Posts and Telecommunications, Chinese Academy of Sciences, University of Chinese Academy of Sciences	China	—

No.	Citing paper	Citing institution(s)	Country	S2
26	3d packing for self-supervised monocular depth estimation	Toyota Research Institute	United States	Methodology
27	Is pseudo-lidar needed for monocular 3d object detection?	Toyota Research Institute	United States	Methodology
28	Depth-aware generative adversarial network for talking head video generation	Alibaba Cloud, Hong Kong University of Science and Technology	China, Hong Kong	Methodology
29	Deep learning sensor fusion for autonomous vehicle perception and localization: A review	Jordan University of Science and Technology, Université Gustave Eiffel, University of British Columbia	Canada, France, Jordan	Methodology
30	Tartanair: A dataset to push the limits of visual slam	Carnegie Mellon University, Microsoft Research, The Chinese University of Hong Kong	China, United States	Background

Showing the 30 most-cited of 846 independent citing papers.

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 Self-supervised visual feature learning with deep neural networks: A survey

"Mehtod SubCategory Code Contribution VideoGAN [85] Generation 3 Forerunner of video generation with GAN MocoGAN [86] Generation 3 Decomposing motion and content for video generation with GAN TemporalGAN [144] Generation 3 Decomposing temporal and image generator for video generation Video Colorization [145] Generation 3 Employing video colorization as the pretext task Un-LSTM [37] Generation 3 Forerunner of video prediction with LSTM ConvLSTM [146] Generation 3 Employing Convolutional LSTM for video prediction MCNet [147] Generation 3 Disentangling motion and content for video prediction LSTMDynamics [148] Generation 7 Learning by predicting long-term temporal dynamic in videos Video Jigsaw [87] Context 7 Learning by jointly reasoning about spatial and temporal context Transitive [31] Context 7 Learning inter and intra instance variations with a Triplet loss 3DRotNet [28] Context 7 Learning by recognizing rotations of video clips CubicPuzzles [27] Context 7 Learning by solving video cubic puzzles ShuffleLearn [40] Context 3 Employing temporal order verification as the pretext task LSTMPermute [149] Context 3 Learning by temporal order verification with LSTM OPN [39] Context 3 Using frame sequence order recognition as the pretext task O3N [29] Context 7 Learning by identifying odd video sequences ArrowTime [90] Context 3 Learning by recognizing the arrow of time in videos TemporalCoherence [150] Context 7 Learning with the temporal coherence of features of frame sequence FlowNet [151] Cross Modal 3 Forerunner of optical flow estimation with ConvNet FlowNet2 [152] Cross Modal 3 Better architecture and better performance on optical flow estimation UnFlow [153] Cross Modal 3 An unsupervised loss for optical flow estimation CrossPixel [23] Cross Modal 7 Learning by predicting motion from a single image as the pretext task CrossModel [24] Cross Modal 7 Optical flow and RGB correspondence verification as pretext task AVTS [25] Cross Modal 7 Visual and Audio correspondence verification as pretext task AudioVisual [26] Cross Modal 3 Jointly modeling visual and audio as fused multisensory representation LookListenLearn [93] Cross Modal 3 Forerunner of Audio-Visual Correspondence for self-supervised learning AmbientSound [154] Cross Modal 7 Predicting a statistical summary of the sound from a video frame EgoMotion [155] Cross Modal 3 Learning by predicting camera motion and the scene structure from videos LearnByMove [94] Cross Modal 3 Learning by predicting the camera transformation from a pairs of images TiedEgoMotion [95] Cross Modal 7 Learning from ego-motor signals and video sequence GoNet [156] Cross Modal 3 Jointly learning monocular depth, optical flow and ego-motion estimation from videos DepthFlow [157] Cross Modal 3 Depth and optical flow learning using cross-task consistency from videos VisualOdometry [158] Cross Modal 3 An unsupervised paradigm for deep visual odometry learning ActivesStereoNet [159] Cross Modal 3 End-to-end self-supervised learning of depth from active stereo systems"

METHODOLOGY Surroundocc: Multi-camera 3d occupancy prediction for autonomous driving

"While early methods require full depth annotations to supervise the depth estimation model [15, 25, 66], later research focuses on self-supervised depth estimation as it does not require intensive human annotations [70, 17, 5, 67, 62, 56, 69, 43, 9]."

METHODOLOGY Deep learning for 3d reconstruction, augmentation, and registration: A review paper

"With a primary focus on inferring depth maps as the scene geometry output, this study has demonstrated success in learning 3D volumetric representations from 2D observations using the concepts of projective geometry [81]."

METHODOLOGY Lite-mono: A lightweight cnn and transformer architecture for self-supervised monocular depth estimation

"Similar to [45] the learning objective is modeled to minimize an image reconstruction loss L r between a target image I t and a synthesized target image ˆ I t , and an edge-aware smoothness loss L smooth constrained on the predicted depth map D t ."

METHODOLOGY Unifying flow, stereo and depth estimation

> *"Monocular methods [65], [66], [67], [68], [69], [70] take a single image as input and use generic network architectures like ResNet [71] to predict the dense depth map, while multi-view methods [12], [20], [37], [72], [73], [74], [75], [76], [77] usually focus on how to encode the geometric inductive bias (cost volume, warping, etc."*

FOLLOW-UP WORK

Stereo magnification: Learning view synthesis using multiplane images

2018 · arXiv preprint arXiv:1805.09817, 2018 · 1,579 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Tapir: Tracking any point with per-frame initialization and temporal refinement	Google DeepMind, University College London, University of Oxford	United Kingdom	—
2	Benchmarking neural radiance fields for autonomous robots: An overview	China Aerodynamics Research and Development Center, City College of New York, Hangzhou Dianzi University	China, United Kingdom, United States	—
3	Recent Trends in 3D Reconstruction of General Non-Rigid Scenes	Google, Max Planck Institute for Informatics, Saarland University	China, Germany, United Kingdom	—
4	NeLF-Pro: neural light field probes for multi-scale novel view synthesis	ETH Zürich, University of Tübingen, Westlake University	China, Germany, Switzerland	—
5	Meshlrm: Large reconstruction model for high-quality meshes	Adobe Research, University of California San Diego	United States	—
6	Advances in feed-forward 3d reconstruction and view synthesis: A survey	Caltech, Harvard University, Harvard University, MIT	China, Germany, Singapore	—
7	Genxd: Generating any 3d and 4d scenes	Microsoft, Microsoft Corporation, National University of Singapore	Singapore, United States	—
8	Freesplatter: Pose-free gaussian splatting for sparse-view 3d reconstruction	Tencent PCG, The University of Hong Kong	Hong Kong	—
9	Vidu4d: Single generated video to high-fidelity 4d reconstruction with dynamic gaussian surfels	Beijing National Research Center for Information Science and Technology (BNRist), Tsinghua University, Beijing Normal University, Tsinghua University	China	—
10	Advances in 4d generation: A survey	University of Nevada Reno, Zhejiang University	China, United States	—
11	Lrm-zero: Training large reconstruction models with synthesized data	Adobe Research, Kiel University, Stony Brook University	Germany, United States	—
12	Baking gaussian splatting into diffusion denoiser for fast and scalable single-stage image-to-3d generation and reconstruction	Adobe Research, Johns Hopkins University, The Hong Kong University of Science and Technology	China, United States	—
13	Sparse-view 3D reconstruction: Recent advances and open challenges	Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences	China	—
14	Nerf: Representing scenes as neural radiance fields for view synthesis	Google, Google Research, UC Berkeley	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
15	Mvsnerf: Fast generalizable radiance field reconstruction from multi-view stereo	Adobe Research, ShanghaiTech University, University of California, San Diego	China, United States	—
16	Vggt: Visual geometry grounded transformer	Meta AI, University of Oxford	United Kingdom, United States	—
17	ShotVerse: Advancing Cinematic Camera Control for Text-Driven Multi-Shot Video Creation	Hong Kong University of Science and Technology, Tencent, The Hong Kong University of Science and Technology	China, Hong Kong	—
18	Nerf in the wild: Neural radiance fields for unconstrained photo collections	Google Brain, Google Research	United States	—
19	Deferred neural rendering: Image synthesis using neural textures	Max Planck Institute for Intelligent Systems, Stanford University, Technical University of Munich	Germany, United States	—
20	Scene representation networks: Continuous 3d-structure-aware neural scene representations	Stanford University	United States	—
21	Neural volumes: Learning dynamic renderable volumes from images	Meta	Israel, United States	—
22	State of the art on diffusion models for visual computing	Adobe Research, Google Research, KAUST	Canada, Germany, Israel	—
23	Wonderland: Navigating 3d scenes from a single image	Snap Inc., University of California, Los Angeles, University of Toronto	Canada, United States	—
24	latentsplat: Autoencoding variational gaussians for fast generalizable 3d reconstruction	Max Planck Institute for Informatics, Saarland University	Germany	—
25	Gram: Generative radiance manifolds for 3d-aware image generation	Microsoft Research, Microsoft Research Asia, Tsinghua University	China	—
26	Neural rays for occlusion-aware image-based rendering	Max Planck Institute for Informatics, Texas A&M University, The University of Hong Kong	China, Germany, Hong Kong	—
27	Graf: Generative radiance fields for 3d-aware image synthesis	Google, Max Planck Society, Universität Tübingen	China, Germany, United States	—
28	Nerf-sr: High quality neural radiance fields using supersampling	Kuaishou Technology, Tsinghua University, University of Pennsylvania	China, United States	—
29	Dust3r: Geometric 3d vision made easy	Aalto University, Naver Labs Europe	Finland, France	—
30	Nope-nerf: Optimising neural radiance field with no pose prior	Mohammed Bin Zayed University of Artificial Intelligence, University of Oxford	United Arab Emirates, United Kingdom	—

Showing the 30 most-cited of 849 independent citing papers.

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

Stereo magnification: learning view synthesis using multiplane images

2018 · 1 citations (GS)

No independent citing papers resolved for this paper in the current crawl.

Contribution 2

Claim — Contribution 2

The researcher pioneered conditional adversarial networks for image-to-image translation, establishing a foundational framework that enabled subsequent advancements in complex, multi-modal generative tasks such as synchronized human motion synthesis.

The researcher's contribution centers on the development of conditional adversarial networks for image-to-image translation, as demonstrated in their seminal 2017 paper. This work serves as the cornerstone for a broader research trajectory that extends into specialized generative applications, including the 2019 follow-up study on synchronized dance generation. This line of work appears to address the challenge of mapping structured inputs to complex visual outputs, moving from general translation tasks to more nuanced, temporally coherent generation problems.

The originality of this approach lies in its application of adversarial learning to conditional generation, a method that seems to have opened new avenues for controlling image synthesis. By building upon the core framework established in 2017, the researcher's later work suggests a progression toward handling more intricate data structures, such as human motion, indicating a sustained effort to refine and expand the capabilities of generative models.

The significance of this research is evidenced by the substantial citation counts, with the core paper accumulating over 31,000 citations and the follow-up work receiving more than 1,100. Furthermore, the vast majority of citations originate from independent researchers, suggesting that this framework has been widely adopted and integrated into the broader scientific community's work on image generation and translation.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1,446 · **183 flagged influential by Semantic Scholar**

CORE PAPER

Image-to-image translation with conditional adversarial networks

2017 · Proceedings of the IEEE conference on computer vision and pattern ..., 2017 · 31,156 citations (GS)

Field-normalised: 22,075 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2017 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Diffusion models: A comprehensive survey of methods and applications	Beijing University of Posts and Telecommunications, Carnegie Mellon University, Peking University	China, United States	Methodology
2	The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges	Northwest Missouri State University	United States	—
3	Data augmentation: A comprehensive survey of modern approaches	Cape Coast Technical University, University of Mines and Technology	Ghana	—

No.	Citing paper	Citing institution(s)	Country	S2
4	Self-supervised visual feature learning with deep neural networks: A survey	City University of New York, The Graduate Center, The City University of New York	United States	Background
5	A comprehensive survey of loss functions and metrics in deep learning	Centro de Investigaciones en Óptica A.C., Instituto Politecnico Nacional, Universidad Autónoma de Querétaro	Mexico	—
6	A survey on diffusion models for inverse problems	Google, KAIST, Sony AI	South Korea, United States	—
7	AI Art and its Impact on Artists	Artist, California State University Northridge, Distributed Artificial Intelligence Research Institute	Canada, United States	**Methodology**
8	Arbitrary style transfer in real-time with adaptive instance normalization	Cornell University	—	**Methodology**
9	Data augmentation techniques in time series domain: a survey and taxonomy	Universidad Complutense de Madrid, Universidad Politécnica de Madrid	Spain	**Methodology**
10	Deep visual domain adaptation: A survey	Beijing University of Posts and Telecommunications	China	**Methodology**
11	Densely connected pyramid dehazing network	Adobe Research, Johns Hopkins University	United States	Background
12	Deepsdf: Learning continuous signed distance functions for shape representation	Facebook Reality Labs, Massachusetts Institute of Technology, Meta	United States	Background
13	Multimodal unsupervised image-to-image translation	Cornell University, NVIDIA	United States	**Methodology**
14	Deep learning for medical image-based cancer diagnosis	Nanjing Normal University, University of Leicester	China, United Kingdom	—
15	Image data augmentation for deep learning: A survey	Nanjing University, Zhejiang University	China	Background
16	Transformers in medical imaging: A survey	Agency for Science, Technology and Research (A*STAR), Inception Institute of Artificial Intelligence, MBZUAI	Australia, Singapore, United Arab Emirates	**Influential**
17	A survey on deep neural network pruning: Taxonomy, comparison, analysis, and recommendations	Harbin Institute of Technology (Shenzhen), The University of Adelaide, University of Adelaide	Australia, China	—
18	A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions	American International University-Bangladesh, Bangladesh University of Business and Technology, International Islamic University Chittagong	Bangladesh	—
19	A survey of multimodal-guided image editing with text-to-image diffusion models	Fudan University, Nanyang Technological University	China, Singapore	—

No.	Citing paper	Citing institution(s)	Country	S2
20	Shifting machine learning for healthcare from development to deployment and from models to data	Stanford University	United States	Background
21	Stargan v2: Diverse image synthesis for multiple domains	EPFL, Naver, NAVER Corp.	France, Switzerland	Background
22	Deepfakes and beyond: A survey of face manipulation and fake detection	Universidad Autonoma de Madrid	Spain	Methodology
23	Image-to-image translation: Methods and applications	Microsoft, University of Science and Technology of China	China	Methodology
24	Dreamface: Progressive generation of animatable 3d faces under text guidance	Huazhong University of Science and Technology, ShanghaiTech University, The University of Hong Kong	China, Hong Kong	Methodology
25	Self-supervised learning: Generative or contrastive	Anhui University, Beijing Institute of Technology, Renmin University of China	China	—
26	Deep learning in mechanical metamaterials: from prediction and generation to inverse design	Nagoya University, National Institute for Materials Science, University of Tsukuba	Japan	—
27	Deep learning and process understanding for data-driven Earth system science	Max Planck Institute for Biogeochemistry	Germany	—
28	Generative AI design for building structures	Tsinghua University	China	—
29	Neural networks and deep learning	Brown University, University of Pennsylvania	United States	—
30	Recent advances and applications of machine learning in solid-state materials science	Friedrich-Schiller-Universität Jena, Martin-Luther-Universität Halle-Wittenberg	Germany	Background

Showing the 30 most-cited of 853 independent citing papers.

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 Diffusion models: A comprehensive survey of methods and applications

"Generative models have been successfully applied to image restoration tasks, including super-resolution, inpainting and translation [15, 48, 59, 95, 126, 162, 171, 260]."

METHODOLOGY AI Art and its Impact on Artists

"With GANs came the first large-scale image generating models, allowing for output sizes of up to 512 × 512 [24, 61, 65]."

METHODOLOGY Arbitrary style transfer in real-time with adaptive instance normalization

"GANs have also been applied to style transfer [31] and cross-domain image generation [50, 3, 23, 38, 37, 25]."

METHODOLOGY Data augmentation techniques in time series domain: a survey and taxonomy

"This approximation was followed in GANs architectures such as [34, 48, 97, 103, 106], where a comparison between networks is possible using the same loss function to evaluate their training."

METHODOLOGY Deep visual domain adaptation: A survey

"Adversarialbased using domain discriminators to encourage domain confusion through an adversarial objective generative models [59, 3, 48] non-generative models [96] [95, 19, 18, 94, 69]"

FOLLOW-UP WORK

Everybody dance now

2019 · Proceedings of the IEEE/CVF international conference on computer vision ..., 2019 · 1,126 citations (GS)

Field-normalised: 838 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2019 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	CharacterShot: Controllable and Consistent 4D Character Animation	Nanyang Technological University, National University of Singapore, Shanghai AI Lab	China, Singapore	—
2	Fda-gan: Flow-based dual attention gan for human pose transfer	Zhejiang University	China	Methodology
3	Sapiens: Foundation for human vision models	Meta	United States	Background
4	Deep learning for deepfakes creation and detection: A survey	Deakin University, Griffith University, Kumoh National Institute of Technology	Australia, France, Ireland	Methodology
5	Emergence of deepfakes and video tampering detection approaches: A survey	University Institute of Engineering and Technology, Panjab University	India	—
6	Continuous-time video generation via learning motion dynamics with neural ode	KAIST, Korea University, Naver	South Korea	Background
7	A review on generative adversarial networks: Algorithms, theory, and applications	Nanyang Technological University, Southeast University, The University of Sydney	Australia, China, Singapore	Methodology
8	Artificial intelligence in the creative industries: a review	University of Bristol	United Kingdom	Methodology
9	Personalized generation in large model era: A survey	Nanyang Technological University, National University of Singapore, University of Chinese Academy of Sciences	China, Singapore, United States	—
10	First order motion model for image animation	Snap Inc., University of Trento	Italy, United States	Background
11	Melgan: Generative adversarial networks for conditional waveform synthesis	Lyrebird AI, MILA, University of Montreal	Canada	Background
12	Media forensics and deepfakes: an overview	Università Federico II di Napoli	Italy	Background
13	Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward: Deepfakes generation and detection: state-of-the-art, open ...	Oakland University, University of Engineering and Technology-Taxila, University of Michigan-Dearborn	Pakistan, United States	Influential
14	Singan: Learning a generative model from a single natural image	Google Research, Technion, Technion – Israel Institute of Technology	Israel	Background
15	Animatable neural radiance fields for modeling dynamic human bodies	University of California, Berkeley, Zhejiang University	China, United States	Background

No.	Citing paper	Citing institution(s)	Country	S2
16	The creation and detection of deepfakes: A survey	Georgia Institute of Technology	United States	Methodology
17	One-shot free-view neural talking-head synthesis for video conferencing	NVIDIA	United States	Background
18	Deepvoxels: Learning persistent 3d feature embeddings	Max Planck Institute for Intelligent Systems, Princeton University, Stanford University	Germany, United States	Background
19	Text2human: Text-driven controllable human image generation	Nanyang Technological University	Singapore	Background
20	Stylerig: Rigging stylegan for 3d control over portrait images	Massachusetts Institute of Technology, Max Planck Institute for Informatics, Stanford University	France, Germany, United States	—
21	Animating arbitrary objects via deep motion transfer	Snap Inc., University of Trento, University of Trento; Fondazione Bruno Kessler	Italy, United States	Background
22	Deepfake generation and detection, a survey	Wuhan University	China	Background
23	Ai-generated content (aigc) for various data modalities: A survey	Lancaster University, Max Planck Institute for Informatics	Germany, United Kingdom	—
24	Controllable person image synthesis with attribute-decomposed gan	Alibaba Group, ByteDance, Peking University	China	—
25	Latent image animator: Learning to animate images via latent space navigation	Inria	France	Background
26	Generative adversarial networks for image and video synthesis: Algorithms and applications	Cornell University, NVIDIA, University of Illinois at Urbana-Champaign	United States	Background
27	Liquid warping gan: A unified framework for human motion imitation, appearance transfer and novel view synthesis	Hong Kong University of Science and Technology, ShanghaiTech University, Tencent	China, Hong Kong	Methodology
28	Joint generative and contrastive learning for unsupervised person re-identification	Inria, The Chinese University of Hong Kong	China, France	Background
29	Dg-font: Deformable generative networks for unsupervised font generation	East China Normal University	China	Methodology
30	Text2Sign: towards sign language production using neural machine translation and generative adversarial networks	University of Surrey	United Kingdom	Methodology

Showing the 30 most-cited of 593 independent citing papers.

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 Fda-gan: Flow-based dual attention gan for human pose transfer

> *"EDN fails to preserve the source pose and adapt the original target joint to the right positions when there is high variance between the source and target, as shown in the Fig."*

METHODOLOGY Deep learning for deepfakes creation and detection: A survey

"verybodyDanceNow - Automatically transfer the motion from a source to a target person by learning a video-to-video translation. - Can create a motion-synchronized dancing video with multiple subjects [53]. Neural Voice Puppetry https://justusthies.github.io/posts/neuralvoice-puppetry - A method for audio-driven facial video synthesis. - Synthesize videos of a talking head from an audio sequence of ano"

METHODOLOGY A review on generative adversarial networks: Algorithms, theory, and applications

"MoCoGan [74] was proposed to decompose motion and content to generate videos [241], [242]."

METHODOLOGY Artificial intelligence in the creative industries: a review

"ent generation [16{18] [19{32] [25,28], TM [33], RL [34{37], for text, audio, video [38{42] BERT [43], VAE [44] and game [30] NEAT [45] Animation [46{50] [51,52] [53] VAE [52] AR/VR [54,55] Deepfakes [56] [57] VAE [58] Content and captions [59] [60{62] [63,64] VAE [59] Analysis Ads/lm analysis [65] [65] [65] GP [66] Content retrieval [67{70] [71] [72] PM [73] Fake detection [74,75] [74] Blockchain [7"

METHODOLOGY The creation and detection of deepfakes: A survey

"[25] Everybody Dance Now:"

Contribution 3

Claim — Contribution 3

The researcher developed foundational methods to mitigate dataset bias, establishing a critical framework for ensuring fairness and robustness in machine learning systems.

The researcher's primary contribution centers on the seminal 2012 paper 'Undoing the Damage of Dataset Bias,' which appears to address the critical challenge of bias inherent in training data. This work stands alone as a core contribution, with no follow-up papers by the same researcher listed in this specific line of inquiry, suggesting it established a distinct and complete theoretical or methodological advance at the time of publication.

The originality of this work lies in its early and direct confrontation with dataset bias, a problem that has since become central to ethical AI. By focusing on 'undoing' this damage, the researcher likely introduced novel techniques or frameworks to correct or account for skewed data distributions, filling a gap in early machine learning literature that often assumed representative datasets.

The significance of this contribution is evidenced by its substantial impact, with 660 citations indicating widespread recognition. Notably, 99.2% of the citing papers originate from independent researchers, demonstrating that the work has been broadly adopted and validated by the wider scientific community rather than relying on self-citation or institutional echo chambers. This high degree of independent uptake underscores the work's foundational role in the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 0

CORE PAPER

Undoing the Damage of Dataset Bias

2012 · European Conference in Computer Vision, 2012 · 660 citations (GS)

Field-normalised: 571 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2012 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	195
Zhejiang University	China	SCImago #6 · THE 39 · QS 49	157
Nanyang Technological University	Singapore	SCImago #137	134
Peking University	China	SCImago #11 · THE 13 · QS 14	131
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	123
Stanford University	United States	SCImago #18 · THE =5 · QS 3	111
Adobe Research	United States	—	111
University of Oxford	United Kingdom	SCImago #26 · THE 1 · QS 4	105
The Chinese University of Hong Kong	Hong Kong	SCImago #163 · THE =41 · QS =32	98
Google Research	United States	—	96
University of Science and Technology of China	China	SCImago #77 · THE 51 · QS =132	90
Shanghai Jiao Tong University	China	SCImago #10 · THE 40 · QS =47	87
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	86
Google	United States	—	83
Tencent	United States	—	83

Geographic distribution of citing authors

Country	Citing papers
China	1,724
United States	1,517
United Kingdom	396
Germany	288
Singapore	234
Australia	199
South Korea	196
Canada	167
Hong Kong	162
Switzerland	153
France	121
Italy	109

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.

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).

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	Unsupervised learning of depth and ego-motion from video	1,695	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Image-to-image translation with conditional adversarial networks	1,446	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	Undoing the Damage of Dataset Bias	0	8 CFR 204.5(i)(3) — Outstanding Researcher