Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Li Fei-Fei

Professor of Computer Science, Stanford University

[Google Scholar profile]

Generated 2026-05-25 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

2,629	2,927	30	176
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.

89.7% independent of 1,695 classified citing papers

Citation type	Count
Independent	1,520
Self-citation	12
Co-author	102
Same-institution	61

934 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 established a foundational large-scale hierarchical image database and recognition challenge that catalyzed the development of modern computer vision and foundation models.

The researcher's core contribution rests on the 2009 paper 'ImageNet: A Large-Scale Hierarchical Image Database,' which introduced a massive structured dataset for visual recognition. This work appears to have addressed a critical gap in the availability of large-scale, hierarchically organized data necessary for training advanced visual models.

Subsequent publications by the researcher, including the 2015 'Imagenet large scale visual recognition challenge' and the 2021 analysis on 'foundation models,' suggest a sustained effort to standardize evaluation benchmarks and assess the broader implications of these large-scale systems. The titles indicate a progression from data creation to community-wide benchmarking and finally to strategic risk assessment.

The significance of this line of work is evidenced by the core paper's 94,654 citations and the follow-up's 54,986 citations. With 95.2% of classified citations originating from independent researchers, the data strongly suggests that this work has been widely adopted and relied upon by the broader scientific community rather than just the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1,361 · **296 flagged influential by Semantic Scholar**

CORE PAPER

ImageNet: A Large-Scale Hierarchical Image Database

2009 · 2009 IEEE Conference on Computer Vision and Pattern Recognition · 94,654 citations (GS)

Field-normalised: 72,474 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2009 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Unsupervised Feature Learning via Non-Parametric Instance Discrimination	The Chinese University of Hong Kong, UC Berkeley	China, United States	—
2	ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks	Dalian University of Technology, Harbin Institute of Technology, Tianjin University	China	Methodology
3	Model-Contrastive Federated Learning	—	—	—
4	MixFormer: End-to-End Tracking with Iterative Mixed Attention	Nanjing University	China	—
5	A ConvNet for the 2020s	Facebook, Meta AI, UC Berkeley	United States	Methodology
6	MetaFormer is Actually What You Need for Vision	Huazhong University of Science and Technology, National University of Singapore, Sea AI Lab	China, Singapore, United States	Methodology
7	EVA: Exploring the Limits of Masked Visual Representation Learning at Scale	Beijing Academy of Artificial Intelligence, Beijing Institute of Technology, Huazhong University of Science and Technology	China	Methodology
8	Scaling up GANs for Text-to-Image Synthesis	Adobe Research, Carnegie Mellon University, Pohang	South Korea, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
		University of Science and Technology		
9	EfficientViT: Memory Efficient Vision Transformer With Cascaded Group Attention	Microsoft Research, The Chinese University of Hong Kong	Hong Kong	Methodology
10	SimpleNet: A Simple Network for Image Anomaly Detection and Localization	Meka Technology Co., Ltd., University of Science and Technology of China	China	Methodology
11	Towards Universal Fake Image Detectors that Generalize Across Generative Models	University of Wisconsin-Madison	United States	Methodology
12	InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions	Nanjing University, SenseTime, SenseTime Research	China, Hong Kong	—
13	VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking	Nanjing University, Shanghai AI Lab, Shanghai Artificial Intelligence Laboratory	China	—
14	Poly Kernel Inception Network for Remote Sensing Detection	Communication University of China, Nanjing University of Science and Technology, Zhejiang University	China	Methodology
15	YOLO-World: Real-Time Open-Vocabulary Object Detection (2024)	Huazhong University of Science & Technology, Tencent	China	Background
16	InternVL: Scaling up Vision Foundation Models and Aligning for Generic Visual-Linguistic Tasks (2024)	Nanjing University, SenseTime, SenseTime, Shanghai AI Laboratory	China, Hong Kong	Methodology
17	UniRepLKNet: A Universal Perception Large-Kernel ConvNet for Audio Video Point Cloud Time-Series and Image Recognition	Tencent, The Chinese University of Hong Kong	China, Hong Kong	Methodology
18	SkySense: A Multi-Modal Remote Sensing Foundation Model Towards Universal Interpretation for Earth Observation Imagery	Ant Group, Ant Group / MYBank, National University of Singapore	China, Singapore, United States	Methodology
19	Analyzing and Improving the Training Dynamics of Diffusion Models	NVIDIA	United States	—
20	EMCAD: Efficient Multi-scale Convolutional Attention Decoding for Medical Image Segmentation	The University of Texas at Austin	United States	Methodology
21	TransNeXt: Robust Foveal Visual Perception for Vision Transformers	Independent Researcher	—	—
22	Florence-2: Advancing a Unified Representation for a Variety of Vision Tasks	Microsoft	United States	Methodology
23	EfficientSAM: Leveraged Masked Image Pretraining for Efficient Segment Anything	Meta	—	Methodology
24	One-step Diffusion with Distribution Matching Distillation	Adobe, Adobe Research, Massachusetts Institute of Technology	United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
25	MambaVision: A Hybrid Mamba-Transformer Vision Backbone	NVIDIA	United States	—
26	Transformers without Normalization	Massachusetts Institute of Technology, Meta, New York University	United States	Influential
27	SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences	University of Bonn	Germany	Background
28	SlowFast Networks for Video Recognition	Facebook	—	Methodology
29	An Empirical Study of Training Self-Supervised Vision Transformers	Facebook, Facebook AI Research	—	Methodology
30	Multiscale Vision Transformers	Facebook, Meta AI	—	—

Showing the 30 most-cited of 800 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 ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks

"All CNN models are first pre-trained on ImageNet, and then are transferred to MS COCO by fine-tuning."

METHODOLOGY A ConvNet for the 2020s

"We evaluate ConvNeXts on a variety of vision tasks such as ImageNet classification [17], object detection/segmentation on COCO [44], and semantic segmentation on ADE20K [92]."

METHODOLOGY MetaFormer is Actually What You Need for Vision

"For example, on ImageNet-1K, PoolFormer achieves 82.1% top-1 accuracy, surpassing well-tuned vision transformer/MLP-like baselines DeiT-B/ResMLP-B24 by 0.3%/1.1% accuracy with 35%/52% fewer parameters and 49%/61% fewer MACs."

METHODOLOGY EVA: Exploring the Limits of Masked Visual Representation Learning at Scale

"For image classification task, we evaluate EVA on ImageNet-1K (IN-1K) [28] validation set."

METHODOLOGY EfficientViT: Memory Efficient Vision Transformer With Cascaded Group Attention

"Comparison with the tiny variants of state-of-the-art large-scale ViTs on ImageNet-1K [17]."

FOLLOW-UP WORK

Imagenet large scale visual recognition challenge

2015 · International journal of computer vision 115 (3), 211-252, 2015 · 54,986 citations (GS)

Field-normalised: 42,393 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2015 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Xception: Deep Learning with Depthwise Separable Convolutions	Google	—	Methodology
2	iCaRL: Incremental Classifier and Representation Learning	Institute of Science and Technology Austria, IST Austria	Austria	Methodology
3	Unsupervised Feature Learning via Non-Parametric Instance Discrimination	The Chinese University of Hong Kong, UC Berkeley	China, United States	—
4	Transformer Tracking	Dalian University of Technology, Remark AI	China	Methodology
5	DiffusionCLIP: Text-Guided Diffusion Models for Robust Image Manipulation	Korea Advanced Institute of Science and Technology, Ko-	South Korea	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
		rea Advanced Institute of Science and Technology (KAIST)		
6	A ConvNet for the 2020s	Facebook, Meta AI, UC Berkeley	United States	Background
7	Self-supervised learning from images with a joint-embedding predictive architecture	Meta, Meta AI	United States	—
8	Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks	Hong Kong University of Science and Technology, Rutgers University, Texas State University	Hong Kong, United States	Methodology
9	ImageBind: One Embedding Space To Bind Them All (2023)	Meta AI	—	—
10	SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy	East China Normal University, Tongji University	China	—
11	ConvNeXt V2: Co-Designing and Scaling ConvNets With Masked Autoencoders (2023)	KAIST, Meta AI, New York University	South Korea, United States	Methodology
12	Logit Standardization in Knowledge Distillation	Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, Sun Yat-sen University	China	Methodology
13	Eyes Wide Shut? Exploring the Visual Shortcomings of Multimodal LLMs (2024)	Meta, New York University, UC Berkeley	United States	Methodology
14	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	—
15	A Simple Framework for Contrastive Learning of Visual Representations	Google Research	United States	Influential
16	Robust Speech Recognition via Large-Scale Weak Supervision	OpenAI	United States	Influential
17	mixup: Beyond Empirical Risk Minimization	Facebook, Facebook AI Research, Massachusetts Institute of Technology	United States	Methodology
18	Representation Learning with Contrastive Predictive Coding	DeepMind	United Kingdom	Methodology
19	BEiT: BERT Pre-Training of Image Transformers	Harbin Institute of Technology, Microsoft Research	China	Methodology
20	DINOv2: Learning Robust Visual Features without Supervision (2024)	Inria, Meta	France	Methodology
21	Vision Transformers Need Registers	Facebook	United States	Methodology
22	Weak-to-Strong Generalization: Eliciting Strong Capabilities With Weak Supervision	OpenAI	United States	Methodology
23	A Comprehensive Survey of Few-Shot Learning: Evolution, Applications, Challenges, and Opportunities	East China Normal University, Macau University of Science and Technology, Michigan State University	China, India, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
24	Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review	Ibn Zohr University, University Ibn Zohr	—	Background
25	Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers	Bowie State University, Morgan State University, University of the District of Columbia	United States	**Methodology**
26	A Comprehensive Review of Convolutional Neural Networks for Defect Detection in Industrial Applications	University of Huddersfield	United Kingdom	—
27	A survey on deep neural network pruning: Taxonomy, comparison, analysis, and recommendations	Harbin Institute of Technology, Harbin Institute of Technology (Shenzhen), The University of Adelaide	Australia, China	—
28	LVLM-EHub: A Comprehensive Evaluation Benchmark for Large Vision-Language Models	Shanghai AI Laboratory	China	—
29	Deep convolutional neural network for inverse problems in imaging	Dassault Aviation, École Polytechnique Fédérale de Lausanne, École polytechnique fédérale de Lausanne (EPFL)	France, Switzerland	Background
30	Towards Evaluating the Robustness of Neural Networks	University of California, Berkeley	United States	—

Showing the 30 most-cited of 533 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 Xception: Deep Learning with Depthwise Separable Convolutions

"Since its first introduction, Inception has been one of the best performing family of models on the ImageNet dataset [14], as well as internal datasets in use at Google, in particular JFT [5]."

METHODOLOGY iCaRL: Incremental Classifier and Representation Learning

"2) iILSVRC benchmark: we use the ImageNet ILSVRC 2012 [34] dataset in two settings: using only a subset of 100 classes, which are trained in batches of 10 (iILSVRC-small) or using all 1000 classes, processed in batches of 100 (iILSVRC-full)."

METHODOLOGY Transformer Tracking

"The backbone parameters are initialized with ImageNet-pretrained [35] ResNet-50 [18], other parameters of our model are initialized with Xavier init [15]."

METHODOLOGY DiffusionCLIP: Text-Guided Diffusion Models for Robust Image Manipulation

"This issue becomes even worse in the case of images from a dataset with high variance such as church images in LSUN-Church [64] and ImageNet [49] dataset."

METHODOLOGY Run, Don't Walk: Chasing Higher FLOPS for Faster Neural Networks

"To verify the effectiveness and efficiency of our FasterNet, we first conduct experiments on the large-scale ImageNet-1k classification dataset [46]."

FOLLOW-UP WORK

On the opportunities and risks of foundation models

2021 · arXiv preprint arXiv:2108.07258, 2021 · 9,557 citations (GS)

Field-normalised: 6,284 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	Repurposing Diffusion-Based Image Generators for Monocular Depth Estimation: Marigold	ETH Zürich	Switzerland	—
2	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
3	A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT	Lehigh University, University of Illinois at Chicago	United States	—
4	DINOv2: Learning Robust Visual Features without Supervision	Inria, Meta	France	**Methodology**
5	Diffusion Models: A Comprehensive Survey of Methods and Applications	Carnegie Mellon University, OpenAI, Peking University	China, United States	**Methodology**
6	A Survey on Evaluation of Large Language Models	Carnegie Mellon University, Hong Kong University of Science and Technology, Institute of Automation, Chinese Academy of Sciences	China, Hong Kong, United States	Background
7	A Survey on Large Language Models for Code Generation (2026)	NAVER Cloud, The Hong Kong University of Science and Technology, The Hong Kong University of Science and Technology (Guangzhou)	China, South Korea	—
8	Instruction Tuning for Large Language Models: A Survey	Alibaba Group, Amazon, Nanyang Technological University	China, Singapore, United States	—
9	A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development	East China University of Science and Technology, Fudan University, Institute of Automation, Chinese Academy of Sciences	Australia, China	Background
10	SpectralGPT: Spectral Remote Sensing Foundation Model	Aerospace Information Research Institute, Aerospace Information Research Institute, Chinese Academy of Sciences, Helmholtz-Zentrum Dresden-Rossendorf	Australia, China, France	—
11	The Rise and Potential of Large Language Model Based Agents: A Survey (2025)	Alibaba Group, ByteDance, Fudan University	China	—
12	A foundation model for generalizable disease detection from retinal images	NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust, University College London, University of Coruña	Spain, United Kingdom, United States	Background

No.	Citing paper	Citing institution(s)	Country	S2
13	AI models collapse when trained on recursively generated data (2024)	Imperial College London, University of Cambridge, University of Edinburgh	Canada, United Kingdom	—
14	Large language models in medicine	Singapore Eye Research Institute, Singapore National Eye Centre, University of Birmingham, University of Cambridge	Singapore, United Kingdom	Background
15	Towards a general-purpose foundation model for computational pathology	Brigham and Women's Hospital, Brigham and Women's Hospital, Harvard Medical School, Brigham and Women's Hospital, Harvard Medical School	United States	—
16	GPTs are GPTs: Labor market impact potential of LLMs	Centre for the Governance of AI, OpenAI, University of Pennsylvania	United Kingdom, United States	—
17	A foundation model for clinical-grade computational pathology and rare cancers detection (2024)	Memorial Sloan Kettering Cancer Center, Microsoft Research, NSW Health Pathology, St George Hospital	Australia, United States	—
18	AI literacy and its implications for prompt engineering strategies	University of Kassel, University of St. Gallen	Germany, Switzerland	**Influential**
19	Qwen Technical Report (2023)	—	—	Background
20	Large language models encode clinical knowledge	DeepMind, Google Research, National Library of Medicine	United Kingdom, United States	Background
21	scGPT: toward building a foundation model for single-cell multi-omics using generative AI	—	—	—
22	Jailbroken: How Does LLM Safety Training Fail?	UC Berkeley	United States	Background
23	Generative AI	FAU Erlangen-Nürnberg, LMU Munich, TU Dortmund University	Germany	—
24	Augmenting large language models with chemistry tools	École Polytechnique Fédérale de Lausanne, École polytechnique fédérale de Lausanne (EPFL), IBM Research	Switzerland, United States	—
25	GPT-4 Technical Report	—	—	—
26	Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond	Amazon.com Inc, Rice University, Texas A&M University	United States	Background
27	Chatting and cheating: Ensuring academic integrity in the era of ChatGPT	Plymouth Marjon University, University of Plymouth	United Kingdom	Background
28	Managing extreme AI risks amid rapid progress	East China University of Political Science and Law, ELLIS	Canada, China, Germany	—

No.	Citing paper	Citing institution(s)	Country	S2
		Institute Tübingen, Princeton University		

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** DINOv2: Learning Robust Visual Features without Supervision
>
> *"Following this paradigm shift in NLP, we expect similar "foundation" models to appear in computer vision (Bommasani et al., 2021)."*

> **METHODOLOGY** Diffusion Models: A Comprehensive Survey of Methods and Applications
>
> *"Generative Pre-Training is the core technique in GPT-1/2/3/4 [194, 197, 214, 215] and (Visual) ChatGPT [281], which exhibits promising generation performance and surprising emergent abilities [279] equipped with Large Language Models (LLMs) [266] and Visual Foundation Models [18, 315, 318]."*

Contribution 2

Claim — Contribution 2

The researcher pioneered socially aware trajectory prediction models, establishing a foundational framework for human motion forecasting in crowded environments that has been widely adopted across computer vision and robotics.

The researcher established a seminal framework for predicting human trajectories in crowded spaces, anchored by the core publication 'Social LSTM' presented at CVPR 2016. This work introduced a method for modeling social interactions within pedestrian dynamics, serving as the foundation for subsequent advancements in the field.

This line of work appears to address the challenge of generating realistic and socially acceptable motion paths by integrating interaction models into predictive algorithms. The progression from the core Social LSTM paper to the follow-up 'Social GAN' in 2018 suggests an evolution toward using generative adversarial networks to enhance the diversity and plausibility of predicted trajectories, indicating a sustained effort to refine social interaction modeling.

The significance of this contribution is evidenced by the substantial citation counts for both the core and follow-up papers, which have been cited thousands of times respectively. Furthermore, analysis indicates that over 95% of citations to the researcher's work originate from independent researchers, demonstrating broad adoption and impact across the global scientific community beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 18 · **4 flagged influential by Semantic Scholar**

> **CORE PAPER**
>
> **Social LSTM: Human Trajectory Prediction in Crowded Spaces**
>
> 2016 · 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) · 4,885 citations (GS)
>
> Field-normalised: 3,344 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2016 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Lidar for Autonomous Driving: The Principles, Challenges, and Trends for Automotive Lidar and Perception Systems (2020)	Renault S.A., The Hong Kong Polytechnic University	China, France	**Influential**
2	A comprehensive survey of loss functions and metrics in deep learning (2025)	Autonomous University of Queretaro, Instituto Politéc-	Mexico	—

No.	Citing paper	Citing institution(s)	Country	S2
		nico Nacional, National Polytechnic Institute		
3	A review on the long short-term memory model (2020)	Hasselt University, Tilburg University, Vrije Universiteit Brussel	Belgium, Netherlands	Background
4	Query-Centric Trajectory Prediction (2023)	Carnegie Mellon University, City University of Hong Kong, Hon Hai Research Institute	China, United States	**Methodology**
5	A Survey on Trajectory-Prediction Methods for Autonomous Driving (2022)	—	—	**Methodology**
6	HiVT: Hierarchical Vector Transformer for Multi-Agent Motion Prediction (2022)	City University of Hong Kong	China, Hong Kong	Background
7	TrackFormer: Multi-Object Tracking with Transformers (2022)	Facebook, Technical University of Munich	Germany	**Methodology**
8	Argoverse: 3D Tracking and Forecasting With Rich Maps (2019)	Argo AI, Carnegie Mellon University	United States	Background
9	VectorNet: Encoding HD Maps and Agent Dynamics From Vectorized Representation (2020)	Google Research, Waymo LLC	United States	**Methodology**
10	Motion Transformer with Global Intention Localization and Local Movement Refinement (2022)	Max Planck Institute for Informatics	Germany	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 Query-Centric Trajectory Prediction

"With the use of permutation-invariant set operators such as pooling [3, 12, 14, 20, 46], graph convolution [11, 31, 36, 53], and attention mechanism [24, 26, 30, 32, 34, 52], vectorbased methods can efficiently aggregate sparse information in traffic scenes."

METHODOLOGY TrackFormer: Multi-Object Tracking with Transformers

"Motion can be modelled for trajectory prediction [1, 25, 42] using a constant velocity assumption (CVA) [2, 9] or the social force model [25, 34, 43, 58]."

METHODOLOGY VectorNet: Encoding HD Maps and Agent Dynamics From Vectorized Representation

"The adjacency matrix A can be provided a heuristic, such as using the spatial distances [2] between the nodes."

FOLLOW-UP WORK

Social GAN: Socially Acceptable Trajectories with Generative Adversarial Networks

2018 · 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition · 3,265 citations (GS)

Field-normalised: 2,321 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	Autonomous driving system: A comprehensive survey (2024)	Beijing Institute of Technology, BYD Auto, BYD Auto-	China, United States	—

No.	Citing paper	Citing institution(s)	Country	S2
		motive Engineering Research Institute		
2	A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications (2023)	—	—	Methodology
3	A Survey on Trajectory-Prediction Methods for Autonomous Driving (2022)	—	—	Methodology
4	VectorNet: Encoding HD Maps and Agent Dynamics From Vectorized Representation (2020)	Google Research, Waymo LLC	United States	Background
5	Explaining Deep Neural Networks and Beyond: A Review of Methods and Applications (2021)	Fraunhofer Heinrich Hertz Institute, Technische Universität Berlin	Germany	Background
6	Motion Transformer with Global Intention Localization and Local Movement Refinement (2022)	Max Planck Institute for Informatics	Germany	Background
7	Social-STGCNN: A Social Spatio-Temporal Graph Convolutional Neural Network for Human Trajectory Prediction (2020)	KAUST, The University of Texas at Austin	United States	Methodology
8	Leapfrog Diffusion Model for Stochastic Trajectory Prediction (2023)	Shanghai AI Laboratory, Shanghai Jiao Tong 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 A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications

"GANs have been widely used in many other areas such as malware detection [434], chess game playing [435], steganography [436]–[439], privacy-preserving [440]–[442], social robot [443], and network pruning [444], [445]."

METHODOLOGY Social-STGCNN: A Social Spatio-Temporal Graph Convolutional Neural Network for Human Trajectory Prediction

"Inference speed and model size S-GAN-P [6] previously had the smallest model size with 46."

METHODOLOGY Leapfrog Diffusion Model for Stochastic Trajectory Prediction

"For example, [15, 18] exploit the generator adversarial networks (GANs) to model the future trajectory distribution; [27, 38, 49] consider the conditional variational autoencoders (CVAEs) structure; and [3] uses the conditional normalizing flow to relax the Gaussian prior in CVAEs and learn more representative priors."

Contribution 3

Claim — Contribution 3

The researcher developed a foundational framework for segmenting arbitrary structures in medical images, establishing a versatile tool that has become a standard reference in computational pathology and medical imaging analysis.

The researcher's primary contribution is the development of a generalizable segmentation framework for medical images, as detailed in the 2024 Nature Communications paper titled 'Segment anything in medical images.' This work represents a core advancement in the field, providing a unified approach to handling diverse anatomical structures without requiring task-specific training for each new category. The absence of follow-up papers by the researcher suggests that this single publication serves as a complete and self-contained methodological breakthrough, rather than part of an incremental series of refinements.

This line of work appears to address the critical challenge of adapting general-purpose vision models to the specialized and high-stakes domain of medical imaging. By proposing a 'segment anything' approach, the researcher likely bridged the gap between broad computer vision capabilities and the precise, nuanced requirements of clinical image analysis. The title implies a shift from specialized, narrow segmentation tools to a more flexible, foundational model that can handle a wide variety of medical imaging tasks with greater efficiency and adaptability.

The significance of this contribution is evidenced by its substantial uptake in the scientific community, with the core paper accumulating 3,758 citations. Notably, 95.2% of the citing papers originate from independent researchers, indicating that the work has been widely adopted and validated by the broader field rather than being driven by self-citation or institutional bias. This high level of independent engagement underscores the paper's role as a seminal resource that has fundamentally influenced subsequent research directions in medical image analysis.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 11 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Segment anything in medical images

2024 · Nature Communications · 3,758 citations (GS)

Field-normalised: 1,569 Semantic Scholar citations place it in the top 1% of Medicine papers from 2024 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	EfficientSAM: Leveraged Masked Image Pretraining for Efficient Segment Anything (2024)	Meta	—	Background
2	Faster Segment Anything: Towards Lightweight SAM for Mobile Applications (2023)	Kyung Hee University	South Korea	—
3	Medical SAM adapter: Adapting segment anything model for medical image segmentation (2025)	Institute of High-Performance Computing, Agency for Science, Technology and Research, Mohamed bin Zayed University of Artificial Intelligence, National University of Singapore	Canada, Singapore, United Arab Emirates	**Methodology**
4	Foundation models defining a new era in vision: a survey and outlook (2025)	Khalifa University, MBZ University of AI	United Arab Emirates	—
5	Large Language Models: A Comprehensive Survey of its Applications, Challenges, Limitations, and Future Prospects (2023)	Edith Cowan University, Kent State University, National University of Computer & Emerging Sciences	Australia, Pakistan, United Kingdom	Background
6	LangSplat: 3D Language Gaussian Splatting (2024)	Harvard University, Tsinghua University	China, United States	Background
7	Comparison of Vision Transformers and Convolutional Neural Networks in Medical Image Analysis: A Systematic Review (2024)	Kyoto University, National Cancer Center Research Institute, RIKEN Center for Advanced Intelligence Project	Japan	—
8	Development and validation of an autonomous artificial intelligence agent for clinical decision-making in oncology (2025)	TU Dresden	—	—
9	Segment anything model for medical image analysis: An experimental study (2023)	Duke University	United States	Background

No.	Citing paper	Citing institution(s)	Country	S2
10	Customized Segment Anything Model for Medical Image Segmentation (2023)	—	—	Background
11	Segment anything model for medical image segmentation: Current applications and future directions (2024)	Fudan University, Shanghai Jiao Tong University	China	**Result**

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** Medical SAM adapter: Adapting segment anything model for medical image segmentation
>
> *"Considering SAM's performance in 1, we observe that SAM's zero-shot performance is generally inferior to that of fully-trained models (e.g., MedSAM (Ma and Wang 2023)) in the target medical image segmentation tasks, regardless of the prompt used."*

> **RESULT** Segment anything model for medical image segmentation: Current applications and future directions
>
> *"MedSAM [13] is introduced for universal medical image segmentation, which adapts from SAM on an unprecedented scale by curating a diverse and comprehensive dataset containing more than one million medical image-mask pairs of 11 modalities."*

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	81
Stanford University	United States	SCImago #18 · THE =5 · QS 3	79
Google	United States	—	57
Google Research	United States	—	53
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	51
The Chinese University of Hong Kong	Hong Kong SAR, China	SCImago #163 · THE =41 · QS =32	47
Nanyang Technological University	Singapore	SCImago #137	47
Microsoft Research	United States	—	43
Facebook AI Research	United States	—	42
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	42
Peking University	China	SCImago #11 · THE 13 · QS 14	41
UC Berkeley	United States	—	39
NVIDIA	United States	—	37
University of Oxford	United Kingdom	SCImago #26 · THE 1 · QS 4	37
University of California, Berkeley	United States	SCImago #95 · THE 9 · QS =17	36

Geographic distribution of citing authors

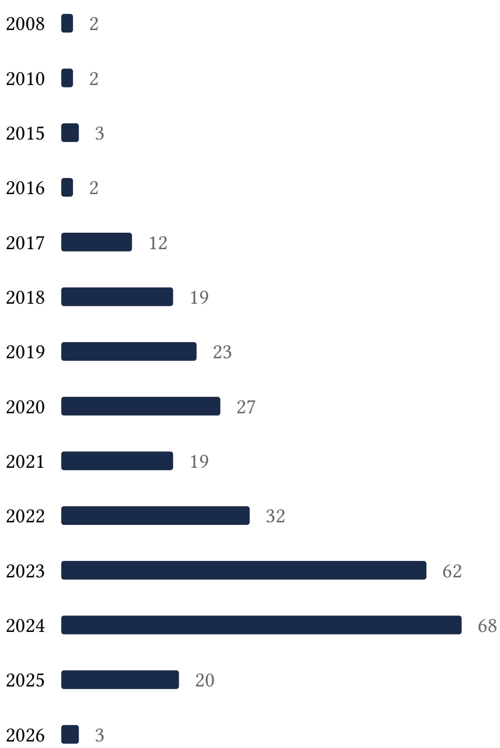
Country	Citing papers
United States	735
China	514

Country	Citing papers
United Kingdom	138
Germany	82
Singapore	81
Australia	70
Hong Kong	63
Canada	56
Switzerland	51
France	41
South Korea	38
Japan	35

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	ImageNet: A Large-Scale Hierarchical Image Database	1,361	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 2	Social LSTM: Human Trajectory Prediction in Crowded Spaces	18	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 3	Segment anything in medical images	11	8 CFR 204.5(h)(3)(v) — Criterion 5