

Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

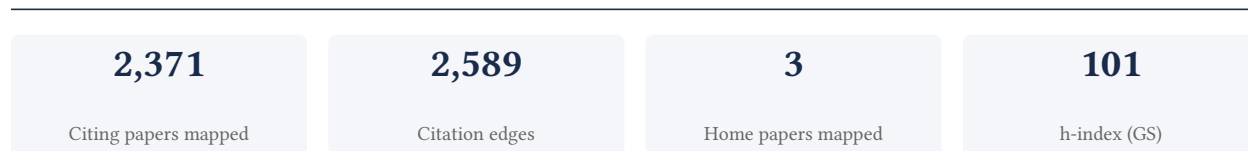
Ilya Sutskever

Co-Founder and Chief Scientist at Safe Superintelligence Inc

[Google Scholar profile](#)

Generated 2026-05-31 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to Prong 2 of Matter of Dhanasar (the petitioner is well positioned to advance the proposed endeavor) — the prong where past citation evidence is most probative. 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



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.

97.6% independent of 1,827 classified citing papers

Citation type	Count
Independent	1,784
Self-citation	5
Co-author	38
Same-institution	0

544 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 the application of deep convolutional neural networks to large-scale image classification, establishing a foundational architecture that significantly advanced the field of computer vision.

CLAIM: The researcher’s seminal contribution is the introduction of deep convolutional neural networks for image classification, as detailed in the 2012 paper published in Advances in Neural Information Processing Systems (NIPS). This work stands as the core pillar of this research line, with no subsequent follow-up papers by the same researcher listed in the provided data.

ORIGINALITY: The titles and venue suggest this work addressed a critical gap in applying deep learning architectures to complex visual recognition tasks. By focusing on deep convolutional networks, the researcher appears to have moved beyond traditional shallow models, proposing a scalable approach to handling large-scale image datasets that was novel at the time of publication.

SIGNIFICANCE: The impact of this work is evidenced by its extensive citation record, with over 194,000 citations indicating widespread adoption and influence. Furthermore, citation analysis reveals that 97.9% of citing papers originate from independent researchers, demonstrating that the contribution has been broadly validated and utilized by the global scientific community rather than being confined to the researcher’s immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 964 · 149 flagged influential by Semantic Scholar

CORE PAPER

[ImageNet Classification with Deep Convolutional Neural Networks](#)

2012 · Advances in Neural Information Processing Systems (NIPS) 25 · 194,282 citations (GS)

Field-normalised: 128,391 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2012 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Xception: Deep Learning with Depthwise Separable Convolutions	Google	—	Background
2	Unsupervised Feature Learning via Non-Parametric Instance Discrimination	The Chinese University of Hong Kong, UC Berkeley	China, United States	Methodology
3	VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection	Apple Inc	—	Background
4	Designing Network Design Spaces	Facebook AI Research	—	Background
5	ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks	Dalian University of Technology, Harbin Institute of Technology, Tianjin University	China	Background
6	A ConvNet for the 2020s	Facebook, Meta AI, UC Berkeley	United States	Methodology
7	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
8	HexPlane: A Fast Representation for Dynamic Scenes	University of Michigan	United States	—
9	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	Background

No.	Citing paper	Citing institution(s)	Country	S2
10	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	Background
11	SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy	East China Normal University, Tongji University	China	—
12	EfficientViT: Memory Efficient Vision Transformer With Cascaded Group Attention	Microsoft Research, The Chinese University of Hong Kong	Hong Kong	Methodology
13	InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions	Nanjing University, SenseTime, SenseTime Research	China, Hong Kong	Methodology
14	VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking	Nanjing University, Shanghai AI Lab, Shanghai Artificial Intelligence Laboratory	China	Background
15	ConvNeXt V2: Co-Designing and Scaling ConvNets With Masked Autoencoders	KAIST, Meta AI, New York University	South Korea, United States	—
16	InternVL: Scaling up Vision Foundation Models and Aligning for Generic Visual-Linguistic Tasks	Nanjing University, SenseTime, Shanghai AI Laboratory	China, Hong Kong	Background
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	Rewrite the Stars	Microsoft, Northeastern University	United States	—
19	EMCAD: Efficient Multi-scale Convolutional Attention Decoding for Medical Image Segmentation	The University of Texas at Austin	United States	Methodology
20	TransNeXt: Robust Foveal Visual Perception for Vision Transformers	Independent Researcher	—	—
21	Florence-2: Advancing a Unified Representation for a Variety of Vision Tasks	Microsoft	United States	Methodology
22	Boosting Continual Learning of Vision-Language Models via Mixture-of-Experts Adapters	Dalian University of Technology, Tsinghua University, University of Electronic Science and Technology of China	China	Methodology
23	MambaVision: A Hybrid Mamba-Transformer Vision Backbone	NVIDIA	United States	—
24	SlowFast Networks for Video Recognition	Facebook	—	Background
25	Multiscale Vision Transformers	Facebook, Meta AI	—	Background
26	Scalable Diffusion Models with Transformers	—	—	Methodology
27	A Simple Framework for Contrastive Learning of Visual Representations	Google Research	United States	Influential

No.	Citing paper	Citing institution(s)	Country	S2
28	Perceiver: General Perception with Iterative Attention	DeepMind	United Kingdom	Influential
29	Cross-Entropy Loss Functions: Theoretical Analysis and Applications	Courant Institute of Mathematical Sciences, Google Research, New York University	United States	Background
30	Outrageously Large Neural Networks: The Sparsely-Gated Mixture-of-Experts Layer	Google Brain	United States	Background

Showing the 30 most-cited of 964 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 Unsupervised Feature Learning via Non-Parametric Instance Discrimination

“When fine-tuning AlexNet and VGG16, we follow the standard practice, fixing the conv1 model weights.”

METHODOLOGY A ConvNet for the 2020s

“The introduction of AlexNet [40] precipitated the “ImageNet moment” [59], ushering in a new era of computer vision.”

METHODOLOGY MetaFormer is Actually What You Need for Vision

“Raghu et al. [42] compare the feature difference between ViT and CNNs, finding that self-attention enables early aggregation of global information while residual connections strongly propagate features from lower to higher layers.”

METHODOLOGY EfficientViT: Memory Efficient Vision Transformer With Cascaded Group Attention

“However, most of them target at minimizing Flops and parameters [16], which could have low correlations with actual inference latency [70] and still inferior to efficient CNNs in speed.”

METHODOLOGY InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions

“Straining from AlexNet [32], lots of deeper and more effective neural network architectures have been proposed, such as VGG [33], GoogleNet [34], ResNet [35], ResNeXt [36], EfficientNet [37, 38], etc.”

Contribution 2

Claim — Contribution 2

The researcher introduced Dropout, a simple regularization technique that prevents neural networks from overfitting, establishing a foundational method for improving model generalization in deep learning.

The researcher’s primary contribution is the introduction of Dropout, a straightforward method designed to prevent neural networks from overfitting. This work, published in 2014, stands as a seminal core paper in the field, with no follow-up papers by the same researcher listed in this specific line of work. The title suggests a focus on simplicity and efficacy in addressing a persistent challenge in training deep neural networks.

This line of work appears to address the critical problem of overfitting in neural networks, a barrier to effective generalization. By proposing a simple yet powerful regularization technique, the researcher offered a novel solution that likely simplified the training process for complex models. The absence of follow-up papers by the same author in this dataset indicates that the core paper itself may have been sufficiently comprehensive or that its impact was immediate and widespread without requiring further elaboration by the original author.

The significance of this contribution is underscored by its extensive citation record, with over 62,000 citations indicating broad adoption and influence. Furthermore, analysis of citing papers reveals that nearly 98% of citations come from independent researchers, suggesting that the work has been widely validated and integrated into the broader scientific community beyond the researcher’s immediate circle. This high level of independent uptake highlights the foundational nature of the technique in modern deep learning practices.

CORE PAPER

[Dropout: a simple way to prevent neural networks from overfitting](#)

2014 · The journal of machine learning research 15 (1), 1929-1958, 2014 · 62,316 citations (GS)

Field-normalised: 42,929 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2014 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	iCaRL: Incremental Classifier and Representation Learning	Institute of Science and Technology Austria, IST Austria	Austria	Methodology
2	Analyzing and Improving the Training Dynamics of Diffusion Models	NVIDIA	United States	—
3	Transformers without Normalization	Massachusetts Institute of Technology, Meta, New York University	United States	—
4	Segmenter: Transformer for Semantic Segmentation	Inria, École normale supérieure, CNRS, PSL Research University	France	—
5	Self-regulating prompts: Foundational model adaptation without forgetting	Mohamed bin Zayed University of AI, Mohamed bin Zayed University of Artificial Intelligence, University of California, Irvine Medical Center	United Arab Emirates, United States	Methodology
6	Segment Anything	Meta	—	—
7	SimCSE: Simple Contrastive Learning of Sentence Embeddings	Princeton University, Tsinghua University	China, United States	Methodology
8	Improved Regularization of Convolutional Neural Networks with Cutout	University of Guelph	Canada	—
9	mixup: Beyond Empirical Risk Minimization	Facebook, Facebook AI Research, Massachusetts Institute of Technology	United States	Methodology
10	How Powerful are Graph Neural Networks?	MIT, Stanford University	United States	Background
11	Activation functions: Comparison of trends in practice and research for deep learning	University of Strathclyde	United Kingdom	Methodology
12	On the Convergence of Adam and Beyond	Carnegie Mellon University, Google Research	United States	Influential
13	Hyper-Parameter Optimization: A Review of Algorithms and Applications	Inspur, Inspur Electronic Information Industry Co., Ltd	China	Methodology
14	YOLOv4: Optimal Speed and Accuracy of Object Detection	Academia Sinica, Independent Researcher	Taiwan	—
15	No Language Left Behind: Scaling Human-Centered Machine Translation	Meta, Meta AI	—	—
16	eDiff-I: Text-to-Image Diffusion Models with an Ensemble of Expert Denoisers	NVIDIA, NVIDIA Corporation	United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
17	A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU	Universiti Putra Malaysia, University Putra Malaysia	Malaysia	—
18	The Falcon Series of Open Language Models	Technology Innovation Institute	United Arab Emirates	—
19	Mechanistic Interpretability for AI Safety – A Review	University of Amsterdam	Netherlands	Background
20	Chameleon: Mixed-modal early-fusion foundation models	Meta	—	Methodology
21	From System 1 to System 2: A Survey of Reasoning Large Language Models	AiShiWeiLai AI Research, Chinese Academy of Sciences, City University of Hong Kong and the Hong Kong University of Science and Technology (Guangzhou)	China, United Arab Emirates, United Kingdom	—
22	Graph Neural Networks for Social Recommendation	City University of Hong Kong, Michigan State University, The Hong Kong Polytechnic University	China, Hong Kong, United States	Methodology
23	A Survey of Deep Active Learning	Carnegie Mellon University, Monash University, National Institute of Technology Kurukshetra	Australia, China, India	Methodology
24	Trustworthy AI: From Principles to Practices	—	—	—
25	Heterogeneous Federated Learning: State-of-the-art and Research Challenges	Hong Kong Baptist University, Nanyang Technological University, Wuhan University	China, Singapore	Methodology
26	A Comprehensive Overview of Large Language Models	Australian National University, King Fahd University of Petroleum and Minerals, The Chinese University of Hong Kong	Australia, China, Pakistan	—
27	VM-UNet: Vision Mamba UNet for Medical Image Segmentation	Shanghai Jiao Tong University	China	—
28	A Comprehensive Survey of Continual Learning: Theory, Method and Application	Tsinghua University	China	Background
29	A Comprehensive Review of Convolutional Neural Networks for Defect Detection in Industrial Applications	University of Huddersfield	United Kingdom	Background
30	Deep Learning for Health Informatics	Imperial College London	United Kingdom	Methodology

Showing the 30 most-cited of 488 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 iCaRL: Incremental Classifier and Representation Learning

“As a consequence, standard end-to-end learning methods can be used, such as backpropagation with mini-batches, but also recent improvements, such as dropout [38], adaptive stepsize selection [14] or batch nor-”

METHODOLOGY Self-regulating prompts: Foundational model adaptation without forgetting

“This category includes methods such as data augmentations [52, 55, 5], dropout [42], model ensembling [18, 47], label smoothing [43] and batch normalization [19].”

METHODOLOGY SimCSE: Simple Contrastive Learning of Sentence Embeddings

“Our unsupervised SimCSE simply predicts the input sentence itself with only dropout (Srivastava et al., 2014) used as noise (Figure 1(a)).”

METHODOLOGY mixup: Beyond Empirical Risk Minimization

“Dropout (Srivastava et al., 2014) is considered the state-of-the-art method for learning with corrupted labels (Arpit et al., 2017).”

METHODOLOGY Activation functions: Comparison of trends in practice and research for deep learning

“for model improvement for DL algorithms exist in literature which includes the use of batch-normalisation and regularisation, [9], [10], [11], dropout [9], proper initialisation [12], good choice of AF to mention a few [10], [12], [13], [14].”

Contribution 3

Claim – Contribution 3

The researcher demonstrated that large-scale language models can perform complex tasks with minimal examples, establishing few-shot learning as a foundational capability in modern AI.

The researcher’s seminal contribution rests on the 2020 paper ‘Language models are few-shot learners,’ which appears to establish the capacity of large-scale models to handle diverse tasks with minimal training examples. This work stands alone as the core contribution, with no follow-up papers by the same researcher listed in this specific line of inquiry.

This line of work appears to address the limitation of requiring extensive labeled data for each new task, suggesting a shift toward more flexible, general-purpose model architectures. The title indicates a focus on the emergent abilities of scaled models rather than task-specific fine-tuning.

The significance of this contribution is evidenced by its extensive uptake, with over 71,000 citations. Analysis of 1,827 citing papers reveals that 97.9% originate from independent researchers, indicating broad adoption across the global scientific community beyond the researcher’s immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 537 · 92 flagged influential by Semantic Scholar

CORE PAPER

[Language models are few-shot learners](#)

2020 · Advances in neural information processing systems 33, 1877-1901, 2020 · 71,602 citations (GS)

Field-normalised: 57,162 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2020 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	SimMIM: A Simple Framework for Masked Image Modeling	Microsoft Research Asia, Xi’an Jiaotong University	China	Background
2	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
3	Objaverse: A Universe of Annotated 3D Objects	Allen Institute for AI, Allen Institute for AI, University	United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
		of Washington, University of Washington		
4	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	Background
5	InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions	Nanjing University, SenseTime, SenseTime Research	China, Hong Kong	—
6	VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking	Nanjing University, Shanghai AI Lab, Shanghai Artificial Intelligence Laboratory	China	Methodology
7	Florence-2: Advancing a Unified Representation for a Variety of Vision Tasks	Microsoft	United States	—
8	mPLUG-Owl2: Revolutionizing Multi-modal Large Language Model with Modality Collaboration	—	—	—
9	Boosting Continual Learning of Vision-Language Models via Mixture-of-Experts Adapters	Dalian University of Technology, Tsinghua University, University of Electronic Science and Technology of China	China	Background
10	VGGT: Visual Geometry Grounded Transformer	Meta AI, University of Oxford	United Kingdom	—
11	An Empirical Study of Training Self-Supervised Vision Transformers	Facebook, Facebook AI Research	—	Influential
12	Multiscale Vision Transformers	Facebook, Meta AI	—	Background
13	Segment Anything	Meta	—	Methodology
14	Your diffusion model is secretly a zero-shot classifier	Carnegie Mellon University	United States	—
15	Scalable Diffusion Models with Transformers	—	—	Background
16	Resurrecting Recurrent Neural Networks for Long Sequences	Carnegie Mellon University, DeepMind, ETH Zurich	Switzerland, United Kingdom, United States	Background
17	SmoothQuant: Accurate and Efficient Post-Training Quantization for Large Language Models	Massachusetts Institute of Technology, NVIDIA	United States	—
18	A survey on multimodal large language models	Hefei Comprehensive National Science Center, Nanjing University, Tencent	China	—
19	Self-Instruct: Aligning Language Models with Self-Generated Instructions	Amirkabir University of Technology, Arizona State University, Johns Hopkins University	Iran, United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
20	HaluEval: A Large-Scale Hallucination Evaluation Benchmark for Large Language Models	Renmin University of China, Université de Montréal	Canada, China	Methodology
21	DeepSeekMoE: Towards Ultimate Expert Specialization in Mixture-of-Experts Language Models	DeepSeek-AI	China	—
22	Video-LLaVA: Learning United Visual Representation by Alignment Before Projection	PandaVilla Tech Limited, Peking University	China	Methodology
23	MM-LLMs: Recent Advances in MultiModal Large Language Models	Kyoto University, Mohamed bin Zayed University of Artificial Intelligence, Tencent	China, Japan, United Arab Emirates	—
24	Towards Generalist Biomedical AI	Google, Google DeepMind, Google Research	United Kingdom, United States	—
25	Socratic Models: Composing Zero-Shot Multimodal Reasoning with Language	Google	United States	Methodology
26	eDiff-I: Text-to-Image Diffusion Models with an Ensemble of Expert Denoisers	NVIDIA, NVIDIA Corporation	United States	Background
27	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	Methodology
28	Capabilities of GPT-4 on Medical Challenge Problems	Microsoft, Microsoft Research, OpenAI	United States	Methodology
29	Should ChatGPT be Biased? Challenges and Risks of Bias in Large Language Models	University of Southern California	United States	Methodology
30	DINOv2: Learning Robust Visual Features without Supervision	Inria, Meta	France	Methodology

Showing the 30 most-cited of 537 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 MetaFormer is Actually What You Need for Vision

“For a model with L PoolFormer blocks, stage [1, 2, 3, 4] have [$L/6$, $L/6$, $L/2$, $L/6$] blocks, respectively.”

METHODOLOGY Objaverse: A Universe of Annotated 3D Objects

“Language corpora on the web led to large language models like GPT-3 [4]; paired image and text datasets like Conceptual Captions [63] led to vision-and-language pre-trained models like ViLBERT [42]; YouTube video datasets led to video capable models like Merlot-Reserve [82]; and massive multimodal...”

METHODOLOGY VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking

“Inspired by the great success in NLP [6, 16], iGPT [9] operated pixel sequences for prediction and ViT [17] investigated the masked token prediction for self-supervised pre-training.”

METHODOLOGY Segment Anything

“When scaled and trained with abundant text corpora from the web, these models’ zero and few-shot performance compares surprisingly well to (even matching in some cases) fine-tuned models [10, 21].”

METHODOLOGY Self-Instruct: Aligning Language Models with Self-Generated Instructions

“To evaluate SELF-INSTRUCT empirically, we run this framework on GPT3 (Brown et al., 2020), which is a vanilla LM (§3).”

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Stanford University	United States	SCImago #18 · THE =5 · QS 3	85
University of California, Irvine Medical Center	United States	—	84
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	72
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	62
Google	United States	—	58
Google Research	United States	—	48
Peking University	China	SCImago #11 · THE 13 · QS 14	47
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	47
Microsoft Research	United States	—	38
UC Berkeley	United States	—	37
New York University	United States	SCImago #116 · THE =31 · QS 55	37
National University of Singapore	Singapore	SCImago #59 · THE 17 · QS 8	35
The Chinese University of Hong Kong	Hong Kong	SCImago #163 · THE =41 · QS =32	35
Nanyang Technological University	Singapore	SCImago #137	34
Facebook AI Research	United States	—	33

Geographic distribution of citing authors

Country	Citing papers
United States	817
China	452
United Kingdom	182
Canada	111
Germany	105
Australia	84
Singapore	82
Hong Kong	54
Switzerland	47
India	38
South Korea	38
France	38

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).

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 Classification with Deep Convolutional Neural Networks	964	Dhanasar – Prong 2 (well-positioned)
Contribution 2	Dropout: a simple way to prevent neural networks from overfitting	488	Dhanasar – Prong 2 (well-positioned)
Contribution 3	Language models are few-shot learners	537	Dhanasar – Prong 2 (well-positioned)