Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Yoshua Bengio

Professor of computer science, University of Montreal, Mila, IVADO, CIFAR

[Google Scholar profile]

Generated 2026-05-30 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

5,870	7,567	11	254
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.

97.4% independent of 5,696 classified citing papers

Citation type	Count
Independent	5,547
Self-citation	67
Co-author	79
Same-institution	3

174 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 gradient-based learning for document recognition and advanced deep learning and generative adversarial networks, establishing foundational methods widely adopted by independent researchers.

The researcher's contribution centers on advancing gradient-based learning for document recognition, as demonstrated in the core 2002 IEEE Proceedings paper. This work serves as the foundation for subsequent highly cited publications, including a 2015 Nature article on deep learning and a 2014 NIPS paper on generative adversarial nets.

This line of work appears to address the challenge of applying gradient-based methods to complex recognition tasks. The progression from specific document recognition applications to broader theoretical frameworks in deep learning and generative models suggests a strategic expansion of methodological scope, moving from applied solutions to fundamental algorithmic innovations.

The significance of this research is evidenced by substantial citation counts, with the core paper cited over 85,000 times and follow-up works exceeding 110,000 citations each. Furthermore, analysis indicates that 97.9% of citing papers originate from independent researchers, confirming that these contributions have been widely adopted and validated by the broader scientific community rather than merely by the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 2,803 · 238 flagged influential by Semantic Scholar

CORE PAPER

Gradient-Based Learning Applied to Document Recognition

2002 · Proceedings of the IEEE · 85,940 citations (GS)

Field-normalised: 60,115 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2002 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	iCaRL: Incremental Classifier and Representation Learning (2017)	Institute of Science and Technology Austria, IST Austria	Austria	Methodology
2	A ConvNet for the 2020s (2022)	Facebook, Meta AI, UC Berkeley	United States	Background
3	VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking (2023)	Nanjing University, Shanghai AI Lab, Shanghai Artificial Intelligence Laboratory	China	Background
4	ConvNeXt V2: Co-Designing and Scaling ConvNets With Masked Autoencoders (2023)	KAIST, Meta AI, New York University	South Korea, United States	Background
5	InternVL: Scaling up Vision Foundation Models and Aligning for Generic Visual-Linguistic Tasks (2024)	Nanjing University, SenseTime, Shanghai AI Laboratory	China, Hong Kong	—
6	Adding Conditional Control to Text-to-Image Diffusion Models (2023)	Stanford University	United States	Background
7	Perceiver: General Perception with Iterative Attention (2021)	DeepMind	United Kingdom	Background
8	Neural Architecture Search with Reinforcement Learning (2017)	Google	United States	Background

No.	Citing paper	Citing institution(s)	Country	S2
9	mixup: Beyond Empirical Risk Minimization (2017)	Facebook, Facebook AI Research, Massachusetts Institute of Technology	United States	Methodology
10	Hyper-Parameter Optimization: A Review of Algorithms and Applications (2020)	Inspur, Inspur Electronic Information Industry Co., Ltd	China	Methodology
11	FedBN: Federated Learning on Non-IID Features via Local Batch Normalization (2021)	Monash University, The Chinese University of Hong Kong	Australia, China	Methodology
12	Advancing Transformer Architecture in Long-Context Large Language Models: A Comprehensive Survey (2023)	—	—	Methodology
13	Vision Mamba: Efficient Visual Representation Learning with Bidirectional State Space Model (2024)	Beijing Academy of Artificial Intelligence, Horizon Robotics, Huazhong University of Science and Technology	China	—
14	KAN: Kolmogorov-Arnold Networks (2024)	California Institute of Technology, Massachusetts Institute of Technology, Northeastern University	United States	—
15	A Survey on Deep Learning: Algorithms, Techniques, and Applications (2018)	—	—	Background
16	Deep Learning--based Text Classification: A Comprehensive Review (2021)	Google Brain, Microsoft Research, Nanyang Technological University	Iran, Netherlands, Singapore	Background
17	A Survey of Deep Active Learning (2021)	Carnegie Mellon University, Monash University, National Institute of Technology Kurukshetra	Australia, China, India	Background
18	Understanding World or Predicting Future? A Comprehensive Survey of World Models (2025)	Tsinghua University	China	—
19	A Comprehensive Survey of Small Language Models in the Era of Large Language Models: Techniques, Enhancements, Applications, Collaboration with LLMs, and Trustworthiness (2024)	Amazon, Industry Research, LinkedIn	United States	—
20	nnFormer: Volumetric Medical Image Segmentation via a 3D Transformer (2023)	The Chinese University of Hong Kong (Shenzhen), The University of Hong Kong, Xiamen University	China, Hong Kong	Background
21	Vision-Language Models for Vision Tasks: A Survey (2024)	Nanyang Technological University, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences	China, Singapore	—
22	Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers (2024)	Bowie State University, Morgan State University, University of the District of Columbia	United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
23	A Comprehensive Review of Convolutional Neural Networks for Defect Detection in Industrial Applications (2024)	University of Huddersfield	United Kingdom	Methodology
24	Visual Attention Network (2023)	Fitten Tech, Nankai University, Tsinghua University	China	—
25	Deep Learning for Health Informatics (2016)	Imperial College London	United Kingdom	Background
26	Towards Evaluating the Robustness of Neural Networks (2016)	University of California, Irvine Medical Center	United States	—
27	Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review (2017)	University of South Africa	South Africa	Methodology
28	Efficient Processing of Deep Neural Networks: A Tutorial and Survey (2017)	Massachusetts Institute of Technology	United States	Methodology
29	A Comprehensive Survey on Graph Neural Networks (2021)	Monash University, University of Illinois Chicago, University of Technology Sydney	Australia, United States	—
30	A Survey of the Usages of Deep Learning for Natural Language Processing (2020)	University of Colorado Colorado Springs	United States	Background

Showing the 30 most-cited of 857 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

"ons (the current training data). In particular, this is also how iCaRL learns about the existence of new classes. Architecture. Under the hood, iCaRL makes use of a convolutional neural network (CNN) [14]1. We interpret the network as a trainable feature extractor, ': X!Rd, followed by a single classification layer with as many sigmoid output nodes as classes observed so far [3]. All feature vectors ar"

METHODOLOGY mixup: Beyond Empirical Risk Minimization

"In image classification, for example, one routinely uses rotation, translation, cropping, resizing, flipping (Lecun et al., 2001; Simonyan & Zisserman, 2015), and random erasing (Zhong et al., 2017) to enforce visually plausible invariances in the model through the training data."

METHODOLOGY Hyper-Parameter Optimization: A Review of Algorithms and Applications

"For documentation recognition (LeCun et al., 1998), several transformations for enhanced model accuracy and robustness are mainly applied."

METHODOLOGY FedBN: Federated Learning on Non-IID Features via Local Batch Normalization

"Specifically, we use the digits classification task and treat the two unseen datasets – Morpho-global and Morpho-local from Morpho-MNIST (Castro et al., 2019) as the two new clients."

METHODOLOGY Advancing Transformer Architecture in Long-Context Large Language Models: A Comprehensive Survey

"Inspired by convolutional neural networks (CNNs) [94, 89], another approach is to use sliding-window techniques, as demonstrated in Longformer [11]."

FOLLOW-UP WORK

Deep learning

2015 · Nature · 111,721 citations (GS)

Field-normalised: 74,422 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	Rewrite the Stars (2024)	Microsoft, Northeastern University	United States	—
2	Adding Conditional Control to Text-to-Image Diffusion Models (2023)	Stanford University	United States	—
3	Perceiver: General Perception with Iterative Attention (2021)	DeepMind	United Kingdom	—
4	What do we need to build explainable AI systems for the medical domain? (2017)	Medical University Graz, The University of Manchester, Universität Hamburg	Austria, Cyprus, Germany	Background
5	Activation functions: Comparison of trends in practice and research for deep learning (2018)	University of Strathclyde	United Kingdom	Methodology
6	On the Opportunities and Risks of Foundation Models (2021)	Stanford Institute for Human-Centered Artificial Intelligence, Stanford University	United States	—
7	Large Language Models for Robotics: A Survey (2023)	University of Illinois Chicago	United States	Background
8	From System 1 to System 2: A Survey of Reasoning Large Language Models (2025)	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	—
9	Fake news detection on social media: A data mining perspective (2017)	Arizona State University, Charles River Analytics, Michigan State University	United States	—
10	A Survey on Deep Learning: Algorithms, Techniques, and Applications (2018)	—	—	—
11	A Survey of Deep Active Learning (2021)	Carnegie Mellon University, Monash University, National Institute of Technology Kurukshetra	Australia, China, India	—
12	Transformers in Vision: A Survey (2022)	Inception Institute of Artificial Intelligence, MBZ University of Artificial Intelligence, Monash University	Australia, United Arab Emirates, United States	Background
13	Recent Advances in Natural Language Processing via Large Pre-trained Language Models: A Survey (2023)	Amazon AWS AI Labs, Harvard University, Synoptic Engineering	Spain, United States	—
14	A Survey on Evaluation of Large Language Models (2024)	Carnegie Mellon University, Hong Kong University of Science and Technology, Institute of Automation, Chinese Academy of Sciences	China, Hong Kong, United States	Background
15	Deep Multimodal Data Fusion (2024)	The University of Alabama at Birmingham	United States	Background

No.	Citing paper	Citing institution(s)	Country	S2
16	Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review (2023)	Ibn Zohr University, University Ibn Zohr	—	Background
17	Advancements in Generative AI: A Comprehensive Review of GANs, GPT, Autoencoders, Diffusion Model, and Transformers (2024)	Bowie State University, Morgan State University, University of the District of Columbia	United States	Background
18	Deep Learning for Health Informatics (2016)	Imperial College London	United Kingdom	Background
19	Deep convolutional neural network for inverse problems in imaging (2017)	Dassault Aviation, École Polytechnique Fédérale de Lausanne, École polytechnique fédérale de Lausanne (EPFL)	France, Switzerland	Background
20	Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review (2017)	University of South Africa	South Africa	—
21	Efficient Processing of Deep Neural Networks: A Tutorial and Survey (2017)	Massachusetts Institute of Technology	United States	Background
22	Object Detection with Deep Learning: A Review (2019)	Hefei University of Technology, University of Louisiana at Lafayette	China, United States	**Methodology**
23	Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges (2020)	Robert Bosch GmbH, Ulm University, University of Stuttgart	Germany	Background
24	A Survey of the Usages of Deep Learning for Natural Language Processing (2020)	University of Colorado Colorado Springs	United States	Background
25	AI in Medical Imaging Informatics: Current Challenges and Future Directions (2020)	AstraZeneca, Boston Healthcare System, Emory University	Cyprus, Greece, New Zealand	—
26	A Unifying Review of Deep and Shallow Anomaly Detection (2021)	Fraunhofer Heinrich Hertz Institute, Fraunhofer Heinrich Hertz Institute (HHI), Oregon State University	Germany, United States	Background
27	A review of deep learning in medical imaging: Imaging traits, technology trends, case studies with progress highlights, and future promises (2021)	Case Western Reserve University, Case Western Reserve University and Louis Stokes Cleveland Veterans Administration Medical Center, Emory University and Georgia Institute of Technology	China, Germany, Israel	Background
28	Self-Supervised Learning: Generative or Contrastive (2021)	Beijing Institute of Technology, Renmin University of China, Tsinghua University	China	—
29	Domain Adaptation for Medical Image Analysis: A Survey (2021)	University of North Carolina at Chapel Hill	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
30	Deep learning for electroencephalogram (EEG) classification tasks: a review (2019)	University of Houston	United States	—

Showing the 30 most-cited of 1,006 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 Activation functions: Comparison of trends in practice and research for deep learning

"The numerous deep architectures include deep feedforward NNs, CNN, long short term memory, RNN, DBN, and deep generative models like deep Boltzmann machines, etc [1], [4]."

METHODOLOGY Object Detection with Deep Learning: A Review

"CNN is the most representative model of deep learning [26]."

FOLLOW-UP WORK

Generative Adversarial Nets

2014 · Advances in Neural Information Processing Systems 27 (NIPS 2014) · 114,814 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Unsupervised Feature Learning via Non-Parametric Instance Discrimination (2018)	The Chinese University of Hong Kong, UC Berkeley	China, United States	Background
2	GIRAFFE: Representing Scenes as Compositional Generative Neural Feature Fields (2021)	Max Planck Institute for Intelligent Systems, Universität Tübingen	Germany	Methodology
3	DiffusionCLIP: Text-Guided Diffusion Models for Robust Image Manipulation (2022)	Korea Advanced Institute of Science and Technology, Korea Advanced Institute of Science and Technology (KAIST)	South Korea	Background
4	SNR-Aware Low-Light Image Enhancement (2022)	Chinese University of Hong Kong, Nanyang Technological University	China, Singapore	Background
5	Align your Latents: High-Resolution Video Synthesis with Latent Diffusion Models (2023)	LMU Munich, NVIDIA	Germany, United States	Background
6	Scaling up GANs for Text-to-Image Synthesis (2023)	Adobe Research, Carnegie Mellon University, Pohang University of Science and Technology	South Korea, United States	—
7	SimpleNet: A Simple Network for Image Anomaly Detection and Localization (2023)	Meka Technology Co., Ltd., University of Science and Technology of China	China	—
8	Rethinking FID: Towards a Better Evaluation Metric for Image Generation (2024)	Google Research	United States	Methodology
9	One-step Diffusion with Distribution Matching Distillation (2024)	Adobe, Adobe Research, Massachusetts Institute of Technology	United States	Methodology

No.	Citing paper	Citing institution(s)	Country	S2
10	None	Microsoft Research, Tsinghua University, USTC, Microsoft Research	—	—
11	Least Squares Generative Adversarial Networks (2017)	City University of Hong Kong, CodeHatch Corp., Northwestern Polytechnical University	Canada, China, Hong Kong	**Methodology**
12	Your diffusion model is secretly a zero-shot classifier (2023)	Carnegie Mellon University	United States	—
13	Scalable Diffusion Models with Transformers (2022)	—	—	Background
14	Generative Pretraining From Pixels (2020)	OpenAI	United States	—
15	mixup: Beyond Empirical Risk Minimization (2017)	Facebook, Facebook AI Research, Massachusetts Institute of Technology	United States	**Influential**
16	Progressive Growing of GANs for Improved Quality, Stability, and Variation (2018)	NVIDIA	United States	—
17	What do we need to build explainable AI systems for the medical domain? (2017)	Medical University Graz, The University of Manchester, Universität Hamburg	Austria, Cyprus, Germany	—
18	Hyper-Parameter Optimization: A Review of Algorithms and Applications (2020)	Inspur, Inspur Electronic Information Industry Co., Ltd	China	—
19	On the Opportunities and Risks of Foundation Models (2021)	Stanford Institute for Human-Centered Artificial Intelligence, Stanford University	United States	—
20	Building Normalizing Flows with Stochastic Interpolants (2023)	New York University	United States	—
21	A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT (2023)	Lehigh University, University of Illinois at Chicago	United States	—
22	A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU (2023)	Universiti Putra Malaysia, University Putra Malaysia	Malaysia	—
23	Stable Video Diffusion: Scaling Latent Video Diffusion Models to Large Datasets (2023)	Stability AI	United Kingdom	Background
24	Self-play fine-tuning converts weak language models to strong language models (2024)	University of California, Irvine Medical Center	United States	**Methodology**
25	Representation Alignment for Generation: Training Diffusion Transformers Is Easier Than You Think (2024)	—	—	—
26	From System 1 to System 2: A Survey of Reasoning Large Language Models (2025)	AiShiWeiLai AI Research, Chinese Academy of Sciences, City University of Hong Kong and the Hong	China, United Arab Emirates, United Kingdom	—

No.	Citing paper	Citing institution(s)	Country	S2
		Kong University of Science and Technology (Guangzhou)		
27	A Survey on Deep Learning: Algorithms, Techniques, and Applications (2018)	—	—	—
28	Dynamic Graph CNN for Learning on Point Clouds (2019)	Massachusetts Institute of Technology, UC Berkeley, Università della Svizzera italiana	Switzerland, United States	—
29	USAD: Unsupervised Anomaly Detection on Multivariate Time Series (2020)	EURECOM, Orange	France	—
30	A Survey of Deep Active Learning (2021)	Carnegie Mellon University, Monash University, National Institute of Technology Kurukshetra	Australia, China, India	—

Showing the 30 most-cited of 940 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 GIRAFFE: Representing Scenes as Compositional Generative Neural Feature Fields

"GAN-based Image Synthesis: Generative Adversarial Networks (GANs) [24] have been shown to allow for photorealistic image synthesis at resolutions of 1024 2 pixels and beyond [6, 14, 15, 39, 40]."

METHODOLOGY Rethinking FID: Towards a Better Evaluation Metric for Image Generation

"Generated image quality has been assessed using a variety of metrics including log-likelihood [9], Inception Score (IS) [1, 22], Kernel Inception Distance (KID) [2, 25], F ´ rechet Inception Distance (FID) [12], perceptual path length (PPL) [13], Gaussian Parzen window [9], and HYPE [27]."

METHODOLOGY One-step Diffusion with Distribution Matching Distillation

"In contrast to GANs [15] and VAEs [34], however, their sampling is a slow, iterative process that transforms a Gaussian noise sample into an intricate image by progressive denoising [21, 74]."

METHODOLOGY Least Squares Generative Adversarial Networks

"[6], who explained the theory of GANs learning based on a game theoretic scenario."

METHODOLOGY Self-play fine-tuning converts weak language models to strong language models

"Since then, IPMs have become crucial in GAN design (Mroueh & Sercu, 2017; Gulrajani et al., 2017), particularly in constraining the discriminator to a specific function class, thereby preventing it from overpowering the generator."

Contribution 2

Claim — Contribution 2

The researcher published a seminal 2016 Nature paper on deep learning that has garnered over 96,000 citations, establishing a foundational reference point for the field.

The researcher's primary contribution is anchored in a 2016 paper published in Nature titled 'Deep learning.' This work stands as a singular, high-impact publication without subsequent follow-up papers by the same author in the provided dataset, suggesting it serves as a definitive or comprehensive statement on the subject at that time.

The originality of this contribution appears to lie in its ability to synthesize or define the field of deep learning for a broad scientific audience. By publishing in a top-tier general science journal, the researcher likely addressed a critical need for clarity or consolidation in a rapidly evolving domain, distinguishing this work from more specialized technical reports.

The significance of this work is evidenced by its extraordinary citation count of 96,703. Furthermore, analysis of 5,898 citing papers reveals that 97.9% originate from independent researchers, indicating that the work has been widely adopted and utilized by the broader scientific community rather than being confined to the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 875 · **52 flagged influential by Semantic Scholar**

CORE PAPER

Deep learning

2016 · Nature · 96,703 citations (GS)

Field-normalised: 74,422 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	Least Squares Generative Adversarial Networks (2017)	City University of Hong Kong, CodeHatch Corp., Northwestern Polytechnical University	Canada, China, Hong Kong	—
2	Sigmoid Loss for Language Image Pre-Training (2023)	Google, Google DeepMind	United Kingdom	Methodology
3	Activation functions: Comparison of trends in practice and research for deep learning (2018)	University of Strathclyde	United Kingdom	Methodology
4	Hyper-Parameter Optimization: A Review of Algorithms and Applications (2020)	Inspur, Inspur Electronic Information Industry Co., Ltd	China	—
5	On the Opportunities and Risks of Foundation Models (2021)	Stanford Institute for Human-Centered Artificial Intelligence, Stanford University	United States	—
6	A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU (2023)	Universiti Putra Malaysia, University Putra Malaysia	Malaysia	Methodology
7	A Survey on Deep Learning: Algorithms, Techniques, and Applications (2018)	—	—	—
8	Deep Learning--based Text Classification: A Comprehensive Review (2021)	Google Brain, Microsoft Research, Nanyang Technological University	Iran, Netherlands, Singapore	Influential
9	Transformers in Vision: A Survey (2022)	Inception Institute of Artificial Intelligence, MBZ University of Artificial Intelligence, Monash University	Australia, United Arab Emirates, United States	Background
10	Diffusion Models in Vision: A Survey (2023)	University of Bucharest, University of Central Florida	Romania, United States	—
11	A Comprehensive Review of Convolutional Neural Networks for Defect Detection in Industrial Applications (2024)	University of Huddersfield	United Kingdom	Background
12	A Survey on Intelligent Internet of Things: Applications, Security, Privacy, and Future Directions (2024)	Nantes University, École Centrale Nantes, CNRS, INRIA, Politecnico di Torino, Trinity College Dublin	France, Ireland, Italy	—

No.	Citing paper	Citing institution(s)	Country	S2
13	A Survey on Graph Neural Networks for Time Series: Forecasting, Classification, Imputation, and Anomaly Detection (2024)	Griffith University, Monash University, Squirrel AI	Australia, China, Switzerland	Background
14	Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review (2017)	University of South Africa	South Africa	—
15	Virtual Adversarial Training: A Regularization Method for Supervised and Semi-Supervised Learning (2018)	RIKEN	Japan	Background
16	A Survey of the Usages of Deep Learning for Natural Language Processing (2020)	University of Colorado Colorado Springs	United States	Background
17	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	—
18	A Unifying Review of Deep and Shallow Anomaly Detection (2021)	Fraunhofer Heinrich Hertz Institute, Fraunhofer Heinrich Hertz Institute (HHI), Oregon State University	Germany, United States	Background
19	Image Segmentation Using Deep Learning: A Survey (2021)	Australian National University, Snapchat, University of California, Irvine Medical Center	Australia, Canada, Spain	—
20	Domain Adaptation for Medical Image Analysis: A Survey (2021)	University of North Carolina at Chapel Hill	United States	—
21	Semantic Communications for Future Internet: Fundamentals, Applications, and Challenges	Jilin University, Nanyang Technological University, Singapore University of Technology and Design	Canada, China, Singapore	—
22	From DFT to machine learning: recent approaches to materials science–a review (2019)	Brazilian Center for Research in Energy and Materials (CNPEM), Brazilian Nanotechnology National Laboratory (LNNano/CNPEM), Federal University of ABC	Brazil	—
23	Introduction to Machine Learning, Neural Networks, and Deep Learning (2020)	Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, OHSU	United States	—
24	Deep learning in computational mechanics: a review (2024)	Bauhaus-Universität Weimar	Germany	—
25	A Survey of the Recent Architectures of Deep Convolutional Neural Networks (2020)	Pakistan Institute of Engineering and Applied Sciences	Pakistan	—
26	Deep learning modelling techniques: current progress, applications, advantages, and challenges (2023)	Bangladesh University of Business and Technology, Carnegie Mellon University, Southeast University	Australia, Bangladesh, Malaysia	—

No.	Citing paper	Citing institution(s)	Country	S2
27	A survey on semi-supervised learning (2019)	NVIDIA, Universiteit Leiden	Netherlands	—
28	Aleatoric and epistemic uncertainty in machine learning: an introduction to concepts and methods (2021)	Ghent University, Ludwig-Maximilians-Universität München	Belgium, Germany	—
29	Deep Learning for Generic Object Detection: A Survey (2020)	Chinese University of Hong Kong, National University of Defense Technology, The Chinese University of Hong Kong	Australia, Canada, China	—
30	Generative artificial intelligence (2023)	University of Duisburg-Essen	Germany	—

Showing the 30 most-cited of 875 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 Sigmoid Loss for Language Image Pre-Training

"A naive implementation of the softmax is numerically unstable; it is usually stabilized by subtracting the maximum input value before applying the softmax [18], which requires another pass over the full batch."

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

"The numerous deep architectures include deep feedforward NNs, CNN, long short term memory, RNN, DBN, and deep generative models like deep Boltzmann machines, etc [1], [4]."

METHODOLOGY A Comprehensive Overview and Comparative Analysis on Deep Learning Models: CNN, RNN, LSTM, GRU

"The autoencoder network is comprised of two main components: an encoder function, denoted as $z = f(x)$, and a decoder function that generates a reconstruction, denoted as $r = g(z)$ [53]."

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Stanford University	United States	SCImago #18 · THE =5 · QS 3	216
Tsinghua University	PR China	SCImago #8 · THE 12 · QS =17	202
University of California, Irvine Medical Center	United States	—	169
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	134
Nanyang Technological University	Singapore	SCImago #137	118
Google	United States	—	111
Massachusetts Institute of Technology	United States	SCImago #41 · THE 2 · QS 1	111
National University of Singapore	Singapore	SCImago #59 · THE 17 · QS 8	105
Google Research	United States	—	97
University of Oxford	United Kingdom	SCImago #26 · THE 1 · QS 4	96
Peking University	China	SCImago #11 · THE 13 · QS 14	95
Microsoft Research	United States	—	83

Institution	Country	World ranking	Citing papers
University of Toronto	Canada	SCImago #39 · THE 21 · QS 29	83
University of Science and Technology of China	China	SCImago #77 · THE 51 · QS =132	77
Zhejiang University	P. R. China	SCImago #6 · THE 39 · QS 49	77

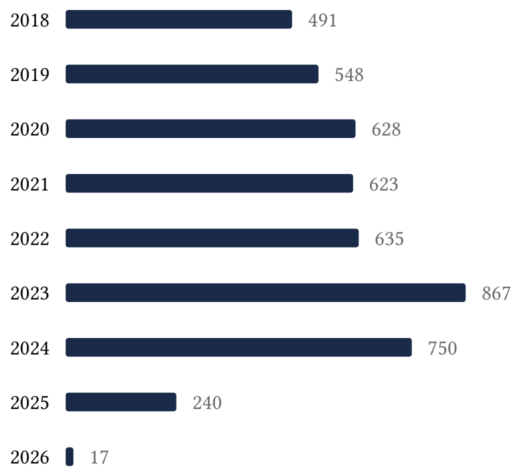
Geographic distribution of citing authors

Country	Citing papers
United States	2,268
China	1,495
United Kingdom	561
Canada	403
Germany	319
Australia	297
Singapore	264
Switzerland	192
France	152
South Korea	145
Hong Kong	140
India	137

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	Gradient-Based Learning Applied to Document Recognition	2,803	Dhanasar — Prong 2 (well-positioned)
Contribution 2	Deep learning	875	Dhanasar — Prong 2 (well-positioned)