Citation Evidence Report

EB-1A Petition — Original Contributions of Major Significance

8 CFR § 204.5(h)(3)(v) · Criterion 5

Kaiming He

Associate Professor, EECS, MIT

Google Scholar profile

Generated 2026-05-31 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

2,719	3,447	96	77
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.

90.9% independent of 1,500 classified citing papers

Citation type	Count
Independent	1,363
Self-citation	27
Co-author	99
Same-institution	11

1,219 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 region proposal networks for real-time object detection, establishing a foundational framework that enabled subsequent breakthroughs in residual learning and instance segmentation.*

The researcher's core contribution rests on the 2015 paper Faster R-CNN, which introduced region proposal networks to advance object detection toward real-time performance. This work serves as the anchor for a broader research line that includes highly cited follow-ups on deep residual learning and Mask R-CNN.

This line of work appears to address the critical challenge of balancing accuracy with computational efficiency in computer vision. The chronological progression from Faster R-CNN to ResNet and Mask R-CNN suggests a systematic effort to refine feature extraction and segmentation capabilities, building upon the initial architectural innovations of the core paper.

The significance of this contribution is evidenced by the massive citation counts for all three papers, indicating widespread adoption. Furthermore, with 94.4% of citing papers originating from independent researchers, the work demonstrates substantial impact beyond the researcher's immediate circle, confirming its role as a foundational resource in the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1,412 · **237 flagged influential by Semantic Scholar**

CORE PAPER

Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks

2015 · Neural Information Processing Systems (NeurIPS), 2015, 2015 · 105,043 citations (GS)

Field-normalised: 71,976 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	Unsupervised Feature Learning via Non-Parametric Instance Discrimination	The Chinese University of Hong Kong, UC Berkeley	China, United States	Methodology
2	VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection	Apple Inc	—	Methodology
3	Video Action Transformer Network	Carnegie Mellon University, DeepMind	United Kingdom, United States	Methodology
4	ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks	Dalian University of Technology, Harbin Institute of Technology, Tianjin University	China	—
5	Transformer Tracking	Dalian University of Technology, Remark AI	China	Methodology
6	Poly Kernel Inception Network for Remote Sensing Detection	Communication University of China, Nanjing University of Science and Technology, Zhejiang University	China	—
7	YOLO-World: Real-Time Open-Vocabulary Object Detection	Huazhong University of Science & Technology, Tencent	China	Background
8	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	—

No.	Citing paper	Citing institution(s)	Country	S2
9	MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI	IN.AI Research, Independent Researcher, Multimodal Art Projection	Canada, India, United States	Methodology
10	Least Squares Generative Adversarial Networks	City University of Hong Kong, CodeHatch Corp., Northwestern Polytechnical University	Canada, China, Hong Kong	Methodology
11	MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications	Google Inc.	—	Methodology
12	YOLOv4: Optimal Speed and Accuracy of Object Detection	Academia Sinica, Independent Researcher	Taiwan	Methodology
13	YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications	Meituan Inc.	—	—
14	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
15	A Survey on Deep Learning: Algorithms, Techniques, and Applications	—	—	Methodology
16	Transformers in Vision: A Survey	Inception Institute of Artificial Intelligence, MBZ University of Artificial Intelligence, Monash University	Australia, United Arab Emirates, United States	—
17	Natural Language Understanding and Inference with MLLM in Visual Question Answering: A Survey	Sun Yat-sen University, Tsinghua University, Worcester Polytechnic Institute	China, United States	—
18	Understanding World or Predicting Future? A Comprehensive Survey of World Models	Tsinghua University	China	—
19	Instruction Tuning for Large Language Models: A Survey	Alibaba Group, Amazon, Nanyang Technological University	China, Singapore, United States	—
20	Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review	Ibn Zohr University, University Ibn Zohr	—	Methodology
21	A Comprehensive Survey of Continual Learning: Theory, Method and Application	Tsinghua University	China	Methodology
22	Vision-Language Models for Vision Tasks: A Survey	Nanyang Technological University, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences	China, Singapore	Influential
23	A Comprehensive Review of Convolutional Neural Networks for Defect Detection in Industrial Applications	University of Huddersfield	United Kingdom	—
24	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	—

No.	Citing paper	Citing institution(s)	Country	S2
25	Object Detection with Deep Learning: A Review	Hefei University of Technology, University of Louisiana at Lafayette	China, United States	—
26	Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges	Robert Bosch GmbH, Ulm University, University of Stuttgart	Germany	**Methodology**
27	A Comprehensive Survey on Graph Neural Networks	Monash University, University of Illinois Chicago, University of Technology Sydney	Australia, United States	—
28	Self-Supervised Visual Feature Learning With Deep Neural Networks: A Survey	City University of New York, The Graduate Center, The City University of New York	United States	**Methodology**
29	A Survey on Performance Metrics for Object-Detection Algorithms	Federal University of Rio de Janeiro	Brazil	**Influential**
30	A Unifying Review of Deep and Shallow Anomaly Detection	Fraunhofer Heinrich Hertz Institute, Fraunhofer Heinrich Hertz Institute (HHI), Oregon State University	Germany, United States	—

Showing the 30 most-cited of 443 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

"We experiment with Fast R-CNN [7] with AlexNet and VGG16 architectures, and Faster R-CNN [32] with ResNet-50."

METHODOLOGY VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection

"HC-baseline also achieves satisfactory performance compared to the state-of-the-art [5], which shows that our base region proposal network (RPN) is effective."

METHODOLOGY Video Action Transformer Network

"CNN [33], NL: Non-local networks [46], P3D: Pseudo-3D convo-"

METHODOLOGY Transformer Tracking

"For most of the popular trackers (such as SiamFC [1], SiamRPN [22], and ATOM [9]), correlation plays a critical role in integrating the template or target information into the regions of interest (ROI)."

METHODOLOGY MMMU: A Massive Multi-discipline Multimodal Understanding and Reasoning Benchmark for Expert AGI

"This work relies on pre-trained visual representations like Faster RCNN features [61] to minimize the training sample complexity."

FOLLOW-UP WORK

Deep Residual Learning for Image Recognition

2016 · 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) · 323,059 citations (GS)

Field-normalised: 226,137 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	Xception: Deep Learning with Depthwise Separable Convolutions	Google	—	Background

No.	Citing paper	Citing institution(s)	Country	S2
2	Unsupervised Feature Learning via Non-Parametric Instance Discrimination	The Chinese University of Hong Kong, UC Berkeley	China, United States	Influential
3	VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection	Apple Inc	—	Background
4	ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks	Dalian University of Technology, Harbin Institute of Technology, Tianjin University	China	Methodology
5	Transformer Tracking	Dalian University of Technology, Remark AI	China	Methodology
6	Model-Contrastive Federated Learning	—	—	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	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	Influential
9	Objaverse: A Universe of Annotated 3D Objects	Allen Institute for AI, Allen Institute for AI, University of Washington, University of Washington	United States	Methodology
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	Influential
11	SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy	East China Normal University, Tongji University	China	Methodology
12	Towards Universal Fake Image Detectors that Generalize Across Generative Models	University of Wisconsin-Madison	United States	Methodology
13	VideoMAE V2: Scaling Video Masked Autoencoders with Dual Masking	Nanjing University, Shanghai AI Lab, Shanghai Artificial Intelligence Laboratory	China	Background
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	—
16	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	Result
17	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

No.	Citing paper	Citing institution(s)	Country	S2
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	Background
20	TransNeXt: Robust Foveal Visual Perception for Vision Transformers	Independent Researcher	—	—
21	Logit Standardization in Knowledge Distillation	Chinese Academy of Sciences, Institute of Information Engineering, Chinese Academy of Sciences, Sun Yat-sen University	China	Background
22	MambaVision: A Hybrid Mamba-Transformer Vision Backbone	NVIDIA	United States	—
23	RandAugment: Practical automated data augmentation with a reduced search space	Google, Google Research	United States	Methodology
24	Least Squares Generative Adversarial Networks	City University of Hong Kong, CodeHatch Corp., Northwestern Polytechnical University	Canada, China, Hong Kong	Methodology
25	Your diffusion model is secretly a zero-shot classifier	Carnegie Mellon University	United States	Methodology
26	A Simple Framework for Contrastive Learning of Visual Representations	Google Research	United States	Influential
27	Perceiver: General Perception with Iterative Attention	DeepMind	United Kingdom	Methodology
28	Learning Transferable Visual Models From Natural Language Supervision	OpenAI	United States	Methodology
29	A survey on multimodal large language models (2024)	Hefei Comprehensive National Science Center, Nanjing University, Tencent	China	—
30	MedCLIP: Contrastive Learning from Unpaired Medical Images and Text	University of Illinois Urbana-Champaign	United States	—

Showing the 30 most-cited of 683 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

"To evaluate our ECA-Net on ImageNet classification, we employ four widely used CNNs as backbone models, including ResNet-50 [11], ResNet-101 [11], ResNet-512 [11] and MobileNetV2 [28]."

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 Model-Contrastive Federated Learning

"For CIFAR-100 and Tiny-Imagenet, we use ResNet-50 [13] as the base encoder."

METHODOLOGY MetaFormer is Actually What You Need for Vision

"Besides, compared with RSB-ResNet ("ResNet Strikes Back") [57] where ResNet [23] is trained with improved training procedure for the same 300 epochs, PoolFormer still performs better."

METHODOLOGY Objaverse: A Universe of Annotated 3D Objects

"We use this strategy to fine-tune the pretrained ResNet-50 Mask-RCNN [25,26] of [3]."

FOLLOW-UP WORK

Mask R-CNN

2017 · 2017 IEEE International Conference on Computer Vision (ICCV) · 51,852 citations (GS)

No.	Citing paper	Citing institution(s)	Country	S2
1	Video Action Transformer Network	Carnegie Mellon University, DeepMind	United Kingdom, United States	—
2	ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks	Dalian University of Technology, Harbin Institute of Technology, Tianjin University	China	—
3	SimMIM: A Simple Framework for Masked Image Modeling	Microsoft Research Asia, Xi'an Jiaotong University	China	—
4	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	—
5	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	—
6	Objaverse: A Universe of Annotated 3D Objects	Allen Institute for AI, Allen Institute for AI, University of Washington, University of Washington	United States	—
7	Poly Kernel Inception Network for Remote Sensing Detection	Communication University of China, Nanjing University of Science and Technology, Zhejiang University	China	—
8	YOLO-World: Real-Time Open-Vocabulary Object Detection	Huazhong University of Science & Technology, Tencent	China	—
9	SpatialVLM: Endowing Vision-Language Models with Spatial Reasoning Capabilities (2024)	Google DeepMind	United Kingdom	—
10	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	—
11	TransNeXt: Robust Foveal Visual Perception for Vision Transformers	Independent Researcher	—	—
12	Florence-2: Advancing a Unified Representation for a Variety of Vision Tasks	Microsoft	United States	—
13	EfficientSAM: Leveraged Masked Image Pretraining for Efficient Segment Anything	Meta	—	—
14	MambaVision: A Hybrid Mamba-Transformer Vision Backbone	NVIDIA	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
15	YOLOv4: Optimal Speed and Accuracy of Object Detection	Academia Sinica, Independent Researcher	Taiwan	—
16	The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision)	Microsoft, University of Washington	United States	—
17	Hallucination of Multimodal Large Language Models: A Survey	National University of Singapore	Singapore	—
18	A Survey on Deep Learning: Algorithms, Techniques, and Applications	—	—	—
19	Transformers in Vision: A Survey	Inception Institute of Artificial Intelligence, MBZ University of Artificial Intelligence, Monash University	Australia, United Arab Emirates, United States	—
20	Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review	Ibn Zohr University, University Ibn Zohr	—	—
21	A Comprehensive Review of Convolutional Neural Networks for Defect Detection in Industrial Applications	University of Huddersfield	United Kingdom	—
22	Visual Attention Network	Fitten Tech, Nankai University, Tsinghua University	China	—
23	Object Detection with Deep Learning: A Review	Hefei University of Technology, University of Louisiana at Lafayette	China, United States	—
24	UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation	Arizona State University	United States	—
25	Deep Multi-Modal Object Detection and Semantic Segmentation for Autonomous Driving: Datasets, Methods, and Challenges	Robert Bosch GmbH, Ulm University, University of Stuttgart	Germany	—
26	Image Segmentation Using Deep Learning: A Survey	Australian National University, Snapchat, University of California, Irvine Medical Center	Australia, Canada, Spain	—
27	Beyond Transmitting Bits: Context, Semantics, and Task-Oriented Communications	Amazon, Imperial College London, Technology Innovation Institute	China, South Korea, Spain	—
28	A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects	Griffith University, Hohai University, The Hong Kong University of Science and Technology	Australia, China, Hong Kong	—
29	Foundation models in robotics: Applications, challenges, and the future	Physical Intelligence, Princeton University, Rhoda AI	Canada, China, United States	—
30	A Survey of the Recent Architectures of Deep Convolutional Neural Networks	Pakistan Institute of Engineering and Applied Sciences	Pakistan	—

Showing the 30 most-cited of 286 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).

Contribution 2

Claim — Contribution 2

The researcher developed Focal Loss, a novel loss function that addresses class imbalance in dense object detection, significantly improving model performance and establishing a new standard in the field.

CLAIM: The researcher's primary contribution is the development of Focal Loss, introduced in the seminal 2017 paper "Focal Loss for Dense Object Detection." This work stands as a foundational piece in the field, with no subsequent follow-up papers by the same researcher listed in this specific line of inquiry, indicating the core paper itself represents the complete and self-contained contribution.

ORIGINALITY: The title suggests the work addresses the specific challenge of class imbalance inherent in dense object detection tasks. By introducing a new loss function, the researcher appears to have identified a critical gap in existing training methodologies, offering a mechanism to down-weight easy examples and focus learning on hard, misclassified instances. This approach represents a methodological innovation aimed at improving the efficiency and accuracy of detection models.

SIGNIFICANCE: The impact of this contribution is evidenced by its extensive citation record, with the core paper accumulating 48,505 citations. Furthermore, analysis of 1,500 citing papers reveals that 94.4% originate from independent researchers, demonstrating that the work has been widely adopted and validated by the broader scientific community rather than just the researcher's immediate circle. This high level of independent uptake underscores the work's fundamental importance and broad utility in advancing object detection research.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 16

CORE PAPER

Focal Loss for Dense Object Detection

2017 · International Conference on Computer Vision (ICCV), 2017, 2017 · 48,505 citations (GS)

Field-normalised: 31,004 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2017 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	SCConv: Spatial and Channel Reconstruction Convolution for Feature Redundancy	East China Normal University, Tongji University	China	—
2	EfficientViT: Memory Efficient Vision Transformer With Cascaded Group Attention	Microsoft Research, The Chinese University of Hong Kong	Hong Kong	—
3	Poly Kernel Inception Network for Remote Sensing Detection	Communication University of China, Nanjing University of Science and Technology, Zhejiang University	China	—
4	YOLO-World: Real-Time Open-Vocabulary Object Detection	Huazhong University of Science & Technology, Tencent	China	Background
5	YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications	Meituan Inc.	—	—
6	A Comprehensive Survey of Continual Learning: Theory, Method and Application	Tsinghua University	China	—
7	End-to-End Autonomous Driving: Challenges and Frontiers	Peking University, Universität Tübingen, University of Tübingen	China, Germany	**Methodology**
8	MobileNetV4: Universal Models for the Mobile Ecosystem	Google	United States	—

No.	Citing paper	Citing institution(s)	Country	S2
9	YOLOv10: Real-Time End-to-End Object Detection	Tsinghua University, University of Sheffield	China, United Kingdom	Background
10	A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS	Instituto Politecnico Nacional, Instituto Politécnico Nacional, Universidad Autónoma de Querétaro	Mexico	Methodology
11	Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection	International Digital Economy Academy, Microsoft Research, Tsinghua University	China	Methodology
12	Structured Adversarial Self-Supervised Learning for Robust Object Detection in Remote Sensing Images	The Hong Kong Polytechnic University	China	Methodology
13	The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection	—	—	—
14	DETRs Beat YOLOs on Real-time Object Detection	Baidu Inc, Peking University	China	Background
15	BiFormer: Vision Transformer with Bi-Level Routing Attention	City University of Hong Kong, SenseTime Research	China	—
16	Large Selective Kernel Network for Remote Sensing Object Detection	Nankai University	China	—

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

Citing-text excerpts — how the field used this work

> **METHODOLOGY** End-to-End Autonomous Driving: Challenges and Frontiers
>
> *"Besides, weighting-based approaches [252], [253] are also commonly used."*

> **METHODOLOGY** A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS
>
> *"To address this issue, their AlignOTA method introduces focal loss [6] into the classification cost and uses the IoU of prediction and ground truth box as the soft label, enabling the selection of aligned samples for each target and solving the problem from a global perspective."*

> **METHODOLOGY** Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection
>
> *"Specifically, we dot product each query with text features to predict logits for each text token and then compute focal loss [28] for each logit."*

> **METHODOLOGY** Structured Adversarial Self-Supervised Learning for Robust Object Detection in Remote Sensing Images
>
> *"It can be observed that the update of θ is achieved by minimizing the detection loss L det . λ represents the weight to balance the localization loss L loc (e.g., smooth-L1 loss [73]) and the classification loss L cls (e.g., focal loss [69])."*

Contribution 3

Claim — Contribution 3

The researcher established the Dark Channel Prior as a foundational method for single image haze removal, creating a seminal framework that has become a standard reference in computational photography.

The researcher's primary contribution is the development of the Dark Channel Prior for single image haze removal, introduced in a 2009 paper. This work stands as a seminal core contribution, with no follow-up papers by the same researcher listed in this specific line of inquiry, indicating the original paper's self-contained impact.

This line of work appears to address the challenge of restoring visibility in hazy images using a single input. The title suggests a novel prior-based approach that likely simplified or improved upon existing methods, establishing a new baseline for visibility enhancement tasks in computer vision.

The significance of this contribution is evidenced by its extensive uptake, with over 11,000 citations. Furthermore, analysis of 1,500 citing papers reveals that 94.4% originate from independent researchers, demonstrating that the work has been widely adopted and validated by the broader scientific community rather than just the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1

CORE PAPER

Single Image Haze Removal using Dark Channel Prior

2009 · Computer Vision and Pattern Recognition (CVPR), 2009, 2009 · 11,866 citations (GS)

Field-normalised: 5,973 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	Atlantis++: Enabling Underwater Depth Estimation with Stable Diffusion and Beyond	Beijing Institute of Technology	China	—

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

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Tsinghua University	China	SCImago #8 · THE 12 · QS =17	67
Facebook AI Research	United States	—	43
Nanyang Technological University	Singapore	SCImago #137	40
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	38
Google Research	United States	—	36
Microsoft Research	United States	—	34
Peking University	China	SCImago #11 · THE 13 · QS 14	34
Google	United States	—	34
Stanford University	United States	SCImago #18 · THE =5 · QS 3	31
National University of Singapore	Singapore	SCImago #59 · THE 17 · QS 8	29
Microsoft	United States	—	27
Zhejiang University	China	SCImago #6 · THE 39 · QS 49	27
University of California, Irvine Medical Center	United States	—	27
The Chinese University of Hong Kong	Hong Kong	SCImago #163 · THE =41 · QS =32	26
NVIDIA	United States	—	26

Geographic distribution of citing authors

Country	Citing papers
United States	565
China	544
United Kingdom	113
Singapore	76
Germany	70
Australia	68
Hong Kong	52
India	50
Canada	49
Switzerland	35
South Korea	33
Japan	27

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	Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks	1,412	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 2	Focal Loss for Dense Object Detection	16	8 CFR 204.5(h)(3)(v) — Criterion 5
Contribution 3	Single Image Haze Removal using Dark Channel Prior	1	8 CFR 204.5(h)(3)(v) — Criterion 5