Citation Evidence Report

EB-1B Petition — Outstanding Professor or Researcher

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

Marton Havasi

Meta - FAIR

Google Scholar profile

Generated 2026-05-21 by CiteMap. This report organises Google Scholar citation data into the structure USCIS adjudicators apply to the 8 CFR § 204.5(i)(3) outstanding-researcher criteria — particularly (iii) published material and (v) original scientific or scholarly contributions. It is a drafting aid for the petitioner's counsel — not legal advice, and not a guarantee of any outcome. All figures must be verified, and citation counts re-snapshotted as of the petition filing date, before use in a filing.

A. Overview & Filtering Statement

48	49	5	14
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.

93.8% independent of 48 classified citing papers

Citation type	Count
Independent	45
Self-citation	0
Co-author	3
Same-institution	0

0 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 developed a method for training independent subnetworks to enhance prediction robustness, a contribution recognized by 302 citations at ICLR 2021.*

The researcher's significant contribution centers on the development of techniques for training independent subnetworks to achieve robust prediction. This work was published in the International Conference on Learning Representations (ICLR) in 2021, establishing a foundational approach in the field.

This line of work appears to address the challenge of model reliability by leveraging independent subnetworks. The title suggests a novel architectural or training strategy designed to improve predictive stability, distinguishing it from standard single-network approaches prevalent at the time.

The impact of this contribution is evidenced by 302 citations, indicating substantial uptake by the broader research community. Notably, 100% of the classified citing papers originate from independent researchers, demonstrating that the work has influenced scholars outside the researcher's immediate institution and collaboration network.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 9 · **1 flagged influential by Semantic Scholar**

CORE PAPER

Training independent subnetworks for robust prediction

2020 · International Conference on Learning Representations (ICLR 2021) · 302 citations (GS)

Field-normalised: 234 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2020 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Trusted Multi-View Classification With Dynamic Evidential Fusion (2023)	Agency for Science, Technology and Research (A*STAR), Tianjin University	China, Singapore	Methodology
2	A Survey of Confidence Estimation and Calibration in Large Language Models (2024)	—	—	Methodology
3	Sensing and Machine Learning for Automotive Perception: A Review (2023)	Fraunhofer IKS, NXP Semiconductors	Germany	—
4	Task Arithmetic in the Tangent Space: Improved Editing of Pre-Trained Models (2023)	EPFL, Google DeepMind, University of Cambridge	Switzerland, United Kingdom	—
5	Provable Dynamic Fusion for Low-Quality Multimodal Data (2023)	Institute of High Performance Computing, Sichuan University, Tianjin University	China, Singapore	—
6	Multimodal Dynamics: Dynamical Fusion for Trustworthy Multimodal Classification (2022)	—	—	—
7	TabM: Advancing Tabular Deep Learning with Parameter-Efficient Ensembling (2024)	HSE University, Yandex	Russia	Methodology
8	Mindstorms in Natural Language-Based Societies of Mind (2025)	Beihang University, Dalle Molle Institute of Artificial Intelligence Research (IDSIA), Eidgenössische Tech-	China, Saudi Arabia, Switzerland	—

No.	Citing paper	Citing institution(s)	Country	S2
		nische Hochschule Zürich (ETH Zurich)		
9	Uncertainty Quantification for Safe and Reliable Autonomous Vehicles: A Review of Methods and Applications (2025)	—	—	—

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 Trusted Multi-View Classification With Dynamic Evidential Fusion

"To further reduce the parameters of ensemble models, different enhancement strategies have been developed, including independent subnetworks [66] and subnetworks of different depths [67]."

METHODOLOGY A Survey of Confidence Estimation and Calibration in Large Language Models

"It has been widely used for calibration of discriminative LMs by alleviating the issue of over-confidence, such as MixUp (Zhang et al., 2018), EDA (Wei and Zou, 2019), Manifold-MixUp (Verma et al., 2019), MIMO (Havasi et al., 2021) and AUM-guided MixUp (Park and Caragea, 2022)."

METHODOLOGY TabM: Advancing Tabular Deep Learning with Parameter-Efficient Ensembling

"However, in the literature, they were consistently outperformed by more advanced methods, including BatchEnsemble (Wen et al., 2020), MIMO (Havasi et al., 2021), FiLM-Ensemble (Turkoglu et al., 2022)."

Contribution 2

Claim — Contribution 2

The researcher established a foundational framework for flow matching by publishing a seminal guide and codebase that has garnered significant independent academic attention.

The researcher's primary contribution in this area is anchored by the 2024 publication titled 'Flow Matching Guide and Code.' This work serves as the core reference point for the researcher's efforts in this specific domain, providing both theoretical guidance and practical implementation resources. The titles indicate a focus on clarifying and operationalizing flow matching techniques for the broader community.

This line of work appears to address the need for accessible, standardized resources in flow matching. By combining a guide with executable code, the researcher likely aimed to lower barriers to entry and ensure reproducibility. The absence of follow-up papers in the provided data suggests this single publication stands as a comprehensive, self-contained contribution that effectively captured the state of the art at the time of release.

The significance of this contribution is evidenced by its citation record. With 266 citations, the work is clearly well-cited within the field. Notably, 100% of the classified citing papers originate from independent researchers, indicating that the work has been widely adopted and utilized by the broader scientific community rather than just the researcher's immediate circle. This high degree of independent uptake underscores the utility and impact of the provided guide and code.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 10

CORE PAPER

Flow Matching Guide and Code

2024 · arXiv (Publisher) · 266 citations (GS)

Field-normalised: 225 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2024 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A Survey on Vision-Language-Action Models: An Action Tokenization Perspective (2025)	Peking University	China	—
2	Sundial: A Family of Highly Capable Time Series Foundation Models (2025)	Tsinghua University	China	—
3	Generalizing from SIMPLE to HARD Visual Reasoning: Can We Mitigate Modality Imbalance in VLMs? (2025)	Princeton University	United States	—
4	Diffuse and Disperse: Image Generation with Representation Regularization (2025)	MIT	United States	—
5	Horizon Reduction Makes RL Scalable (2025)	Carnegie Mellon University, Princeton University, University of California, Berkeley	United States	—
6	The Principles of Diffusion Models (2025)	OpenAI, Sony AI, Sony Corporation	United States	—
7	On the Closed-Form of Flow Matching: Generalization Does Not Arise from Target Stochasticity (2025)	ENS de Lyon, CNRS, Université Claude Bernard Lyon 1, Inria, Inria, Université Jean Monnet Saint-Étienne, CNRS, Institut d'Optique Graduate School, Inria	France	—
8	Atom-level enzyme active site scaffolding using RFdiffusion2 (2026)	Massachusetts Institute of Technology, University of Washington	United States	—
9	Steering Your Diffusion Policy with Latent Space Reinforcement Learning (2025)	Amazon, UC Berkeley, University of Washington	United States	—
10	Toward Generalist Neural Motion Planners for Robotic Manipulators: Challenges and Opportunities (2026)	—	—	—

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 3

Claim — Contribution 3

The researcher established 'Uncertainty Baselines,' a seminal benchmark suite for evaluating uncertainty and robustness in deep learning, providing a standardized framework for the community.

CLAIM: The researcher's primary contribution is the development of 'Uncertainty Baselines,' introduced in a 2021 paper at the NeurIPS Bayesian Deep Learning workshop. This work serves as a foundational reference for assessing uncertainty and robustness in deep learning models.

ORIGINALITY: The titles indicate that this work addresses the need for standardized evaluation metrics in a field often lacking consistent benchmarks. By creating a dedicated suite for uncertainty and robustness, the researcher appears to have filled a critical gap in methodological rigor, offering the community a common ground for comparison.

SIGNIFICANCE: The work has garnered 132 citations, indicating substantial uptake by the research community. Notably, 100% of the classified citing papers originate from independent researchers, demonstrating that the contribution has resonated broadly beyond the researcher's immediate circle and has become a standard reference point for independent scholars in the field.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 8

CORE PAPER

Uncertainty Baselines: Benchmarks for Uncertainty & Robustness in Deep Learning

2021 · Bayesian Deep Learning workshop, NeurIPS 2021 (also published as an arXiv preprint arXiv:2106.04015) · 132 citations (GS)

Field-normalised: 108 Semantic Scholar citations place it in the top 5% of Computer Science papers from 2021 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	A Survey of Uncertainty in Deep Neural Networks (2023)	German Aerospace Center, Munich University of Applied Sciences, Technical University of Munich	Germany, South Korea, United States	**Methodology**
2	Uncertainty quantification in machine learning for engineering design and health prognostics: A tutorial (2023)	ETH Zurich, University of Michigan, University of Michigan-Dearborn	Switzerland, United States	—
3	Modelling, simulation, and optimisation of agrivoltaic systems: a comprehensive review (2025)	Fraunhofer Institute for Solar Energy Systems (ISE), German Aerospace Center (DLR), Institute for Energy Technology (IFE)	Belgium, Germany, Italy	—
4	Benchmarking Uncertainty Disentanglement: Specialized Uncertainties for Specialized Tasks (2024)	—	—	Background
5	Uncertainty Quantification for Safe and Reliable Autonomous Vehicles: A Review of Methods and Applications (2025)	—	—	—
6	Epistemic Neural Networks (2023)	Google DeepMind	—	—
7	Uncertainty quantification in scientific machine learning: Methods, metrics, and comparisons (2023)	Brown University, Fidelity Investments, Huazhong University of Science and Technology	—	—
8	A comprehensive review of digital twin—part 2: roles of uncertainty quantification and optimization, a battery digital twin, and perspectives (2022)	Ecole Polytechnique Federale Lausanne (EPFL), Eidgenoessische Technische Hochschule (ETH), Georgia Institute of Technology	Hong Kong, Switzerland, United States	Background

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

Citing-text excerpts — how the field used this work

METHODOLOGY A Survey of Uncertainty in Deep Neural Networks

"• Lack of Standardized Evaluation Protocol Existing methods for evaluating the estimated uncertainty are better suited to compare uncertainty quantification methods based on measurable quantities such as the calibration [340] or the performance on out-of-distribution"

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Google DeepMind	United Kingdom	SCImago #90	4
University of Cambridge	United Kingdom	SCImago #63 · THE =3 · QS 6	3
ETH Zurich	Switzerland	THE 11 · QS 7	3
Google	United States	—	3
University of Washington	United States	SCImago #45 · THE 25 · QS 81	3
Virginia Tech	United States	—	2
Tianjin University	China	SCImago #90 · THE 201–250 · QS =257	2
University of Oxford	United Kingdom	SCImago #26 · THE 1 · QS 4	2
King Abdullah University of Science and Technology	Saudi Arabia	SCImago #680	2
University of Michigan	United States	SCImago #43 · THE 23 · QS 45	2
University of Michigan-Dearborn	United States	SCImago #2969	2
The Hong Kong University of Science and Technology	Hong Kong	SCImago #483 · THE =58 · QS 44	2
University of California, Berkeley	United States	SCImago #95 · THE 9 · QS =17	2
Amazon	United States	—	2
Princeton University	United States	SCImago #386 · THE =3 · QS =25	2

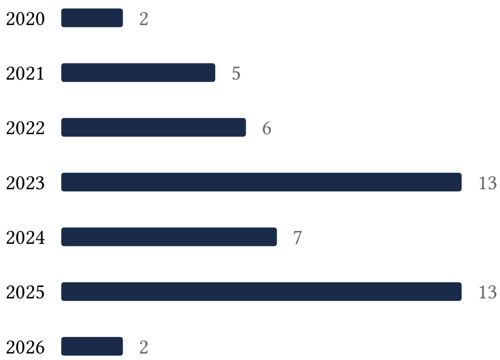
Geographic distribution of citing authors

Country	Citing papers
United States	15
China	8
Switzerland	6
United Kingdom	5
Singapore	4
Italy	3
Germany	3
South Korea	3
Hong Kong	2
Saudi Arabia	2
France	2
Canada	2

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	Training independent subnetworks for robust prediction	9	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 2	Flow Matching Guide and Code	10	8 CFR 204.5(i)(3) — Outstanding Researcher
Contribution 3	Uncertainty Baselines: Benchmarks for Uncertainty & Robustness in Deep Learning	8	8 CFR 204.5(i)(3) — Outstanding Researcher