# Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)


## Hao Tan

Adobe Research

Google Scholar profile


**Generated 2026-05-21 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 | 5 | 4 | 32 |
|---|---|---|---|
| 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.

**80.0% independent** of 5 classified citing papers

| Citation type | Count |
|---|---|
| Independent | 4 |
| Self-citation | 1 |
| Co-author | 0 |
| 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 advanced vision-language navigation by introducing back translation with environmental dropout, a method subsequently extended to evaluate CLIP's utility in broader vision-and-language tasks.*

The researcher established a foundational approach to navigating unseen environments through back translation with environmental dropout, as detailed in a 2019 NAACL paper. This core work serves as the basis for subsequent research, including a 2021 ICLR study examining the benefits of CLIP for vision-and-language tasks.

This line of work appears to address the challenge of generalizing navigation systems to novel settings. The progression from specific navigation techniques to broader evaluations of pre-trained models like CLIP suggests an effort to integrate robust environmental understanding with large-scale vision-language representations.

The impact of this research is evidenced by substantial citation counts, with the core paper cited 463 times and the follow-up work cited 615 times. Furthermore, analysis indicates that 80% of citing papers originate from independent researchers, demonstrating broad adoption and influence across the academic community beyond the researcher's immediate circle.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1**

**CORE PAPER**

### Learning to Navigate Unseen Environments: Back Translation with Environmental Dropout

2019 · North American Chapter of the Association for Computational Linguistics · 463 citations (GS)

Field-normalised: 391 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2019 indexed by Semantic Scholar, by citation count.

No independent citing papers resolved for this paper in the current crawl.

**FOLLOW-UP WORK**

### How Much Can CLIP Benefit Vision-and-Language Tasks?

2021 · International Conference on Learning Representations · 615 citations (GS)

Field-normalised: 501 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2021 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | GPT-4V(ision) is a Generalist Web Agent, if Grounded (2024) | — | — | — |

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

## Contribution 2

### Claim — Contribution 2

*The researcher developed LXMERT, a cross-modality encoder framework that integrates visual and textual representations using transformer architectures, establishing a foundational approach for joint vision-language understanding.*

The researcher's primary contribution centers on the development of LXMERT, introduced in a 2019 paper published at EMNLP-IJCNLP. This work proposes a method for learning cross-modality encoder representations from transformers, aiming to unify visual and linguistic data processing within a single architectural framework. By leveraging transformer-based mechanisms, the research addresses the challenge of effectively aligning and integrating heterogeneous modalities, a critical gap in multimodal artificial intelligence systems at the time.

The originality of this line of work lies in its systematic approach to cross-modal representation learning. Rather than treating vision and language as separate streams, the titles suggest a unified encoder design that captures interactions between modalities. This architectural innovation appears to have provided a scalable and effective baseline for subsequent research in joint vision-language tasks, distinguishing itself through its reliance on transformer technology for cross-modal alignment.

The significance of this contribution is evidenced by its substantial citation count of 3,680, indicating widespread adoption and influence within the natural language processing and computer vision communities. Furthermore, citation analysis reveals that 80% of the citing papers originate from independent researchers, underscoring the work's broad impact beyond the researcher's immediate institution or collaboration network. This high level of independent uptake confirms the framework's utility as a standard reference point in the field.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 2**

**CORE PAPER**

### LXMERT: Learning Cross-Modality Encoder Representations from Transformers

2019 · Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) · 3,680 citations (GS)

Field-normalised: 2,883 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2019 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision) (2023) | Microsoft, University of Washington | United States | — |
| 2 | Deep Multimodal Data Fusion (2024) | The University of Alabama at Birmingham | United States | — |

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

## Contribution 3

### Claim — Contribution 3

*The researcher proposed a unified text-generation framework for vision-and-language tasks, establishing a foundational approach that has garnered significant independent scholarly attention.*

The researcher's core contribution is the development of a unified framework for vision-and-language tasks via text generation, as detailed in their 2021 paper published at the International Conference on Machine Learning. This work stands as a seminal piece in the field, addressing the fragmentation of multimodal AI by proposing a single generative paradigm. The titles suggest this approach consolidates diverse tasks under one methodological umbrella, offering a novel structural solution to complex cross-modal challenges. The significance of this contribution is evidenced by its substantial citation count of 737, indicating widespread adoption and influence within the academic community. Furthermore, the high proportion of independent citations suggests that the work has resonated beyond the researcher's immediate circle, validating its broad impact and utility for other scholars in the domain.

**INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 1**

**CORE PAPER**

### [Unifying Vision-and-Language Tasks via Text Generation]

2021 · International Conference on Machine Learning · 737 citations (GS)

Field-normalised: 620 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2021 indexed by Semantic Scholar, by citation count.

| No. | Citing paper | Citing institution(s) | Country | S2 |
|---|---|---|---|---|
| 1 | Parameter-Efficient Fine-Tuning for Large Models: A Comprehensive Survey (2024) | Arizona State University, City University of New York (Baruch College & Graduate Center), New York University | Canada, United States | — |

Independent citing papers only; self- and co-author citations excluded. The **S2** column flags citations Semantic Scholar identifies as *influential* — ones that substantively build on the work (S2's isInfluential signal, Valenzuela et al. 2015) — the "built on / relied upon" pattern the AAO credits. Counsel should quote the citing text for the strongest of these.

## D. Citing-Institution Prestige & Geography

### Top citing institutions

| Institution | Country | World ranking | Citing papers |
|---|---|---|---|
| University of Toronto | Canada | SCImago #39 · THE 21 · QS 29 | 1 |
| The University of Alabama at Birmingham | United States | QS 1001-1200 | 1 |
| University of Washington | United States | SCImago #45 · THE 25 · QS 81 | 1 |
| Northeastern University | United States | QS 384 | 1 |
| Microsoft | United States | — | 1 |
| Arizona State University | United States | SCImago #357 · THE 201–250 · QS =173 | 1 |
| University of California, Santa Cruz | United States | SCImago #1349 · THE =181 · QS =458 | 1 |
| New York University | United States | SCImago #116 · THE =31 · QS 55 | 1 |
| University of California, Riverside | United States | SCImago #949 · THE 301–350 · QS =440 | 1 |
| City University of New York (Baruch College & Graduate Center) | United States | SCImago #912 · QS =613 | 1 |

### Geographic distribution of citing authors

| Country | Citing papers |
|---|---|
| United States | 3 |
| Canada | 1 |

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 | Learning to Navigate Unseen Environments: Back Translation with Environmental Dropout | 1 | Dhanasar — Prong 2 (well-positioned) |
| Contribution 2 | LXMERT: Learning Cross-Modality Encoder Representations from Transformers | 2 | Dhanasar — Prong 2 (well-positioned) |
| Contribution 3 | Unifying Vision-and-Language Tasks via Text Generation | 1 | Dhanasar — Prong 2 (well-positioned) |