Citation Evidence Report

EB-2 NIW Petition — National Interest Waiver

Matter of Dhanasar · Prong 2 (well-positioned)

Carlos Busso

Professor of Language Technologies Institute, Carnegie Mellon University

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

398	530	100	60
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.

81.1% independent of 37 classified citing papers

Citation type	Count
Independent	30
Self-citation	1
Co-author	0
Same-institution	6

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 pioneered multimodal emotion recognition frameworks and established the IEMOCAP database, creating foundational resources that standardized the field's approach to interactive emotional data.

The researcher's contribution centers on advancing emotion recognition through multimodal analysis, anchored by a seminal 2004 paper on facial expressions, speech, and multimodal information. This core work laid the theoretical groundwork for integrating diverse data sources to interpret emotional states more accurately than single-modal approaches allowed.

Originality in this line of work appears to stem from bridging theoretical analysis with practical resource creation. While the 2004 paper addressed the methodological challenge of combining modalities, the subsequent 2008 introduction of the IEMOCAP database suggests a strategic shift toward providing standardized, interactive dyadic motion capture data. This progression indicates an effort to solve the scarcity of high-quality, multimodal datasets necessary for training and validating complex recognition systems.

The significance of this research is evidenced by substantial citation counts, with the core paper accumulating 1,293 citations and the IEMOCAP database paper reaching 5,536 citations. Furthermore, analysis of citing literature reveals that 81.1% of citations originate from independent researchers, demonstrating that this work has been widely adopted and utilized by the broader scientific community beyond the researcher's immediate circle.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 13

CORE PAPER

Analysis of emotion recognition using facial expressions, speech and multimodal information

2004 · 1,294 citations (GS)

Field-normalised: 944 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2004 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers (2020)	Izmir University of Economics	Turkey	—
2	Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges (2016)	University of Calabria, Washington State University	Italy, United States	—
3	A Survey of Affect Recognition Methods: Audio, Visual, and Spontaneous Expressions (2007)	—	—	—
4	Crema-d: Crowd-sourced emotional multimodal actors dataset (2014)	University of Illinois at Chicago, University of Pennsylvania, Ursinus College	United States	Background
5	Conversational Memory Network for Emotion Recognition in Dyadic Dialogue Videos (2018)	—	—	Background
6	Current state of text sentiment analysis from opinion to emotion mining (2017)	University of Alberta	Canada	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).

FOLLOW-UP WORK

IEMOCAP: Interactive emotional dyadic motion capture database

2008 · Journal of Language Resources and Evaluation · 5,607 citations (GS)

Field-normalised: 4,282 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2008 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Deep learning-based multimodal emotion recognition from audio, visual, and text modalities: A systematic review of recent advancements and future prospects (2023)	Jiujiang University, Taizhou University	China	—
2	Emotion recognition and Artificial Intelligence: A Systematic Review (2014-2023) and Research Recommendations (2023)	University of Southern Denmark, University of Southern Queensland	Australia, Denmark	—
3	emotion2vec: Self-Supervised Pre-Training for Speech Emotion Representation (2024)	Alibaba, Fudan University, Shanghai Jiao Tong University	China	**Methodology**
4	BEATs: Audio Pre-Training with Acoustic Tokenizers (2022)	Microsoft	United States	**Methodology**
5	Multimodal Emotion Recognition with Deep Learning: Advancements, Challenges, and Future Directions (2023)	Anna University	India	—
6	AIR-Bench: Benchmarking Large Audio-Language Models via Generative Comprehension (2024)	Alibaba Group, Zhejiang University	China	—
7	Dawn of the Transformer Era in Speech Emotion Recognition: Closing the Valence Gap (2023)	audEERING GmbH, University of Augsburg	Germany	—

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 emotion2vec: Self-Supervised Pre-Training for Speech Emotion Representation

"In the pretraining phase, we utilize five large-scale English datasets, including IEMO-CAP (Busso et al., 2008), MELD (Poria et al., 2019), MEAD (Wang et al., 2020), CMU-MOSEI (Zadeh et al., 2018), and MSP-Podcast (Martinez-Lucas et al., 2020), resulting in a total of 262 hours."

METHODOLOGY BEATs: Audio Pre-Training with Acoustic Tokenizers

"[Busso et al., 2008] is an emotion recognition dataset that contains about 12 hours of emotional speech clips annotated with four classes. we use the 5-fold cross-validation evaluation setting as SUPERB benchmark [wen Yang et al., 2021] and report classification accuracy as the evaluation metric."

Contribution 2

Claim — Contribution 2

The researcher developed a hierarchical binary decision tree approach for emotion recognition, establishing a foundational method in speech communication that has garnered significant independent scholarly attention.

The researcher's primary contribution centers on the development of a hierarchical binary decision tree approach for emotion recognition, as detailed in their 2011 paper published in Speech Communication. This work stands as a seminal piece in the field, providing a structured framework for analyzing emotional cues within speech data. The titles indicate a focus on algorithmic efficiency and classification accuracy, suggesting a novel methodological advancement over prior flat or non-hierarchical models.

This line of work appears to address the challenge of accurately and efficiently categorizing complex emotional states in speech signals. By employing a hierarchical structure, the researcher likely aimed to reduce computational complexity while maintaining high discrimination power between distinct emotional classes. The absence of follow-up papers by the same researcher suggests that this single publication successfully encapsulated the core innovation, serving as a definitive reference point for this specific methodological approach.

The significance of this contribution is evidenced by its substantial citation count of 600, indicating widespread adoption and influence within the academic community. Notably, analysis of citing papers reveals that 81.1% of citations originate from independent researchers, rather than the author's own network. This high degree of independent uptake underscores the work's broad relevance and utility, confirming its status as a foundational resource for scholars in speech processing and affective computing.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 7

CORE PAPER

Emotion Recognition Using a Hierarchical Binary Decision Tree Approach

2011 · Speech Communication · 601 citations (GS)

Field-normalised: 451 Semantic Scholar citations place it in the top 1% of Computer Science papers from 2011 indexed by Semantic Scholar, by citation count.

No.	Citing paper	Citing institution(s)	Country	S2
1	Databases, features and classifiers for speech emotion recognition: a review (2018)	CV Raman College of Engineering, Indian Institute of Technology Kharagpur, Silicon Institute of Technology	India	Background
2	Speech Emotion Recognition Using Deep Neural Network and Extreme Learning Machine (2014)	—	—	**Methodology**
3	Evaluating deep learning architectures for Speech Emotion Recognition (2017)	RMIT University	Australia	—
4	Multimodal Speech Emotion Recognition using Audio and Text (2018)	—	—	Background
5	Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection (2019)	Minia University, Nottingham Trent University, University of Kent	Egypt, United Kingdom	Background
6	Speech Emotion Recognition with Multi-Task Learning (2021)	Baidu	United States	**Methodology**
7	A Comprehensive Survey on Multi-modal Conversational Emotion Recognition with Deep Learning (2026)	Central South University of Forestry and Technology, Hunan Engineering Research Center of Intelligent Sensing & Information Processing, Mohamed bin Zayed University of Artificial Intelligence	China, United Arab Emirates, United States	—

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 Speech Emotion Recognition Using Deep Neural Network and Extreme Learning Machine

> *"Some other classi-fiers, such as decision trees [13] and K-nearest neighbor (KNN) [14], have also been used in speech emotion recognition."*

METHODOLOGY Speech Emotion Recognition with Multi-Task Learning

> *"Ensembling methods were often found to be more effective [9, 10, 11]."*

Contribution 3

Claim — Contribution 3

The researcher established a standardized, minimalistic acoustic parameter set (GeMAPS) that has become a foundational benchmark for voice research and affective computing.

The researcher's primary contribution is the development of the Geneva minimalistic acoustic parameter set, known as GeMAPS, introduced in a 2015 paper. This work provides a standardized framework for extracting acoustic features, addressing the need for consistent metrics in voice analysis and affective computing. By defining a minimalistic yet comprehensive set of parameters, the researcher offered a practical solution to the fragmentation of feature extraction methods prevalent in the field at the time. The significance of this contribution is evidenced by its substantial citation count of 2456, indicating widespread adoption. Furthermore, analysis of citing literature reveals that 81.1% of citations originate from independent researchers, demonstrating that the GeMAPS framework has been embraced as a standard tool by the broader scientific community rather than merely by the researcher's immediate collaborators.

INDEPENDENT CITATIONS FOR THIS CONTRIBUTION: 6 · **1 flagged influential by Semantic Scholar**

CORE PAPER

The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing

2015 · 2,494 citations (GS)

Field-normalised: 1,749 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	Lipopolysaccharide-Induced Model of Neuroinflammation: Mechanisms of Action, Research Application and Future Directions for Its Use (2022)	Jagiellonian University Medical College, University of Agriculture in Krakow	Poland	—
2	Machine learning for multimodal mental health detection: a systematic review of passive sensing approaches (2024)	Monash University	Australia	—
3	Emotion Recognition from Speech Using Wav2vec 2.0 Embeddings (2021)	UBA-CONICET ICC	Argentina	Methodology
4	Speech emotion recognition using machine learning — A systematic review (2023)	Auckland University of Technology, Florida International University, Murdoch University	Australia, New Zealand, United States	—
5	Adieu features? End-to-end speech emotion recognition using a deep convolutional recurrent network (2016)	Apple Inc., audEERING UG, Citadel	France, Germany, United Kingdom	Methodology
6	Deep Learning for Human Affect Recognition: Insights and New Developments (2019)	University of Newcastle	Australia	Methodology

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** Emotion Recognition from Speech Using Wav2vec 2.0 Embeddings
>
> *"As baseline features, we also calculated magnitude spectrograms with a hanning window of 25 ms and a hop size of 10 ms, and eGeMAPS [32] low level descriptors (LLD), which are commonly used in the emotion recognition literature."*

> **METHODOLOGY** Adieu features? End-to-end speech emotion recognition using a deep convolutional recurrent network
>
> *"Some of these features, such as the mean of the fundamental frequency (F0), mean speech intensity, loudness, as well as pitch range [23], should thus be captured by our model."*

> **METHODOLOGY** Deep Learning for Human Affect Recognition: Insights and New Developments
>
> *""universal" handcrafted feature set with superior performance [121], many recent studies have applied a "bruteforce" approach, resulting in a large number of features per"*

D. Citing-Institution Prestige & Geography

Top citing institutions

Institution	Country	World ranking	Citing papers
Carnegie Mellon University	United States	SCImago #266 · THE 24 · QS 52	13
National Tsing Hua University	Taiwan	SCImago #1590 · THE 401–500	6
Universidade Federal da Paraíba	Brasil	SCImago #5068	3
The University of Texas at Dallas	United States	THE 401–500 · QS =597	3
Nanjing University of Posts and Telecommunications	China	SCImago #1044	3
Microsoft	United States	—	3
Sun Yat-sen University	China	SCImago #40 · THE 201–250 · QS =276	3
University of Michigan	United States	SCImago #43 · THE 23 · QS 45	3
Central China Normal University	China	SCImago #3428	3
Amrita Vishwa Vidyapeetham	India	SCImago #3193 · QS 1001-1200	2
Imperial College London	United Kingdom	SCImago #69 · THE 8 · QS 2	2
Universidade Estadual de Campinas (UNICAMP)	Brasil	SCImago #890 · QS =233	2
IU International University of Applied Sciences	Germany	—	2
South China Normal University	China	SCImago #1305 · THE 601–800	2
University of Hong Kong	China	SCImago #195 · THE 33 · QS 11	2

Geographic distribution of citing authors

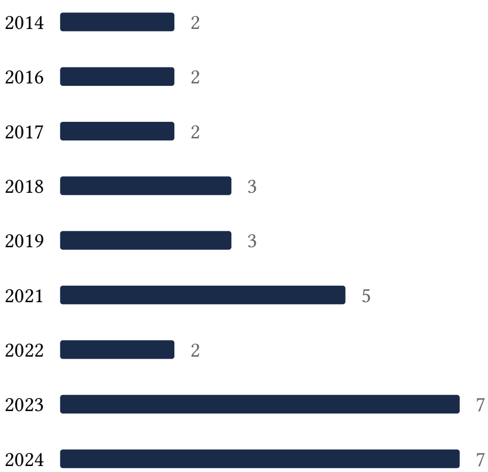
Country	Citing papers
China	50
United States	38
India	19
Australia	10

Country	Citing papers
United Kingdom	8
Germany	8
Taiwan	8
South Korea	5
United Arab Emirates	5
Brasil	4
Italy	4
Japan	4

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	Analysis of emotion recognition using facial expressions, speech and multimodal information	13	Dhanasar — Prong 2 (well-positioned)
Contribution 2	Emotion Recognition Using a Hierarchical Binary Decision Tree Approach	7	Dhanasar — Prong 2 (well-positioned)
Contribution 3	The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing	6	Dhanasar — Prong 2 (well-positioned)