International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Intelligent Trademark Monitoring Using AI: Detection and Prediction of Brand Dilution Risks

Raja Khanna[1], Dr. Chander Parkash Singh[2]

[1]LL.M. Student, University Institute of Legal Studies, Chandigarh University, Mohali, Punjab, India.
[2]Associate Professor, University Institute of Legal Studies, Chandigarh University, Mohali, Punjab, India.

Abstract

Intelligent trademark monitoring combines the recognition of brand identifiers with the assessment of how their use might affect distinctiveness, consumer associations and commercial reputation. These are related but different tasks: detecting a similar logo does not establish dilution, and predicting an adverse brand outcome does not determine legal liability. This review synthesises research on trademark image retrieval, language processing, multimodal interpretation, counterfeit markets, reputation effects and reliable machine learning. It distinguishes directly relevant trademark studies from methods whose application to brand protection remains a proposed transfer. The literature supports image–text retrieval, transformation-aware logo representations and contextual analysis as components of monitoring. Economic and consumer research also shows why risk cannot be inferred from similarity or mention frequency alone: effects depend on product quality, consumer interpretation and the source of reputational change. Recent work on language-model validation, temporal leakage and adaptive production systems provides important safeguards for prediction. The central conclusion is that AI is most defensible as an evidence-producing and prioritisation system, with separate outputs for similarity, contextual concern, future harm and legal review. Progress requires validating the links between these outputs rather than treating a detection benchmark as proof of dilution prevention.

1. Introduction and scope

Trademark monitoring poses a measurement problem as much as a recognition problem. An early multimodal trademark-retrieval report by Sai Ravela and R. Manmatha demonstrated searches using text and visual appearance, including a scalability demonstration involving 63,000 trademark images (Ravela & Manmatha, 2005). That establishes the feasibility of searching substantial design collections, not the legal or commercial significance of the retrieved matches. The review therefore asks: **how far can AI move from detecting brand-related material to predicting a defensible dilution risk?**

The economic motivation extends beyond ownership of an image or name. Ardill-Walker's analysis explains trademarks as signals of otherwise unobservable product variety and quality, reducing consumer search costs while supporting incentives to maintain quality; it also recognises possible barriers to entry (Ardill-Walker, 1994). Monitoring should consequently distinguish protection of informative brand associations from indiscriminate suppression of similar signs.

The discussion follows the monitoring process: defining the outcome, detecting identifiers, interpreting context, predicting change, and evaluating decisions. “Brand dilution” is treated as an interdisciplinary

research subject, not as a synonym for every infringement, counterfeit sale or negative comment. An integrated architecture is presented as a synthesis of the literature, not as an already validated product.

2. Literature-search approach

This narrative review used Semantic Scholar, OpenAlex and PubMed on 12 November 2026. Searches covered trademark monitoring and similarity, logo retrieval, dilution and confusion, counterfeit markets, sentiment analysis, reputation forecasting, and reliable multimodal learning. Broad thematic searches were supplemented by exact-word searches for dilution terminology and searches restricted to 2021–2026 and 2024–2026 for recent developments. The review cites 26 distinct works, selected for their contribution to the monitoring process rather than combined into a quantitative meta-analysis.

The evidence is restricted to dissertations and reports within computer science, economics, education, engineering, environmental science and law; confirmed retractions were excluded. This restriction prevents a comprehensive account of journal and conference landmarks. The cited findings were read from abstracts and retained excerpts, not whole papers, so detailed experimental protocols and independent quality appraisal remain unverified. The legal sources provide historical comparative scholarship rather than an updated statement of Indian or international law. Accordingly, technical and economic findings are separated from proposed applications and legal conclusions.

3. Defining the risk before choosing the model

3.1. Similarity, confusion and dilution are different targets

The legal scholarship retrieved examines different questions. A 2010 dissertation analyses likelihood-of-confusion standards through German, Spanish and UK Community trade mark case law (陸大瑜, 2010). Luo Yue's 2016 dissertation instead examines China's well-known-mark institutions and argues for caution in applying dilution doctrine in that setting (Yue, 2016). Together, these sources warn against translating a numerical resemblance score directly into a universal legal outcome.

For system design, this review proposes separate operational targets:

- **Identifier resemblance:** whether names, sounds or designs are sufficiently similar to warrant inspection.
- **Source-confusion concern:** whether the surrounding presentation suggests an association with a particular commercial source.
- **Distinctiveness erosion:** whether third-party uses appear to weaken the identifier's specific association with the protected brand—the monitoring target corresponding to blurring.
- **Reputational association:** whether a third-party use connects the brand with potentially damaging material—the monitoring target corresponding to tarnishment.

These are working measurement categories, not a complete statutory test. Each requires its own annotation instructions and evidence. A model trained to find copied logos should not have its output relabelled “dilution probability”.

3.2. Consumer meaning is not captured by visual distance alone

Katherine Spencer's questionnaire research found that trademark type affected comprehension but did not significantly affect emotional impact or recall; subject-content compatibility and visual metaphor were associated with higher comprehension (Spencer, 2011). This distinction matters because a representation useful for finding graphical neighbours need not reproduce how consumers understand or remember a mark.

Marketing dilution also needs separation from third-party misuse. Xiaoyan Liu's dissertation on vertical brand extensions found differences in extension evaluations associated with consumers' power-distance beliefs, but no significant difference between high- and low-power-distance consumers in the downward-extension brand-dilution effect (X. Liu, 2016). Its outcome concerns an owner's extension strategy, not a legal determination about another trader. The synthesis is therefore construct-specific: consumer associations can inform monitoring, but the reason those associations change must remain visible.

4. AI detection: from archived designs to real-world use

4.1. Development from handcrafted retrieval to learned representations

The retrieved literature illustrates a progression in technical priorities. Ravela and Manmatha's report, indexed in 2005, combined text search with distributions of curvature and phase derived from image filtering; it reported evaluation on 2,000 trademark images alongside the larger scalability demonstration (Ravela & Manmatha, 2005). Its contribution was multimodal candidate retrieval: the user could start with text and then refine the search visually.

Enjal Parajuli's 2020 thesis moves towards learned representations of logos within larger images. It addresses obfuscations such as colour shifts and modified images, reporting qualitative improvements from an enhanced autoencoder over a general autoencoder and existing content-based retrieval approaches (Parajuli, 2020). The relevant advance is transformation-aware similarity rather than a demonstrated reduction in legal dilution.

A 2021 dissertation on semantic hashing tackles computational cost and memory through compact representations, with models addressing noisy graph construction, semantic preservation and multi-view features (L. Liu, 2021). These contributions are complementary: representation learning seeks meaningful neighbours, while indexing makes retrieving those neighbours economical. A sensible architecture would use compact search for candidate generation and a more discriminating comparison stage for shortlisted pairs; that combination is a design inference, not a benchmark result from these studies.

4.2. Localising logos and reading embedded words

Joshua Song's 2020 thesis uses registered logo images to generate synthetic training data through rendering and domain randomisation. Its CNN detector outperformed previous detectors trained on synthetic data, and the pipeline was demonstrated in robotic object fetching without manually labelled real-world photographs (Song, 2020). This offers directly relevant evidence that trademark collections can supply recognition training material, although the deployment task was robotics rather than enforcement.

Words embedded in packaging require another stage. Zichuan Liu's scene-text dissertation distinguishes text detection from recognition and develops methods for arbitrary shapes and orientations, addressing limitations associated with receptive fields, geometric mismatch and rigid text representations (Z. Liu, 2020). For monitoring, the proposed implication is to preserve both the original crop and the recognised string: a mistaken transcription should not erase the visual evidence.

The studies consequently support a modular detection process rather than one universal classifier. Locating a mark, retrieving similar designs and transcribing lettering answer different questions. Their outputs should remain separately inspectable so that an analyst can distinguish a failed crop from an ambiguous match.

4.3. Unseen marks and changing presentation

A fixed brand vocabulary creates an open-world problem. Xin Sun's 2021 dissertation evaluates open-set recognition on six standard image datasets and industrial pressure data, seeking to classify known inputs while rejecting unknown ones (Sun, 2021). It supports the methodological need for an "unknown" outcome, but does not validate unknown-brand detection.

Zhi Chen's 2023 zero-shot-learning research examines the gap between semantic and visual spaces, including generative reliability, fine-grained generation and shifts between synthetic and real unseen-class distributions (Chen, 2023). Compared with open-set rejection, zero-shot recognition attempts to identify an unseen class using transferable information. These approaches should not be conflated: rejecting unfamiliar imagery and naming a new brand require different evaluations.

Recent robustness work also qualifies architectural optimism. Yiming Bian's 2024 dissertation reports that transformers were generally more robust within its image-classification benchmark and that capacity mattered (Bian, 2024). Jiaxin Qi's 2023 research, however, focuses on how contextual and other spurious correlations can undermine out-of-distribution generalisation (Qi, 2023). The combined judgement is that stronger architectures are plausible candidates, but trademark-specific robustness must be tested against altered packaging, backgrounds and unseen designs rather than assumed from generic benchmarks.

5. Contextual monitoring: what does the detected use mean?

5.1. Text, phonetics and entity resolution

Name matching requires more than exact spelling. Ranjan Satapathy's research examines microtext normalisation, phonetic representations and sentiment analysis; one reported baseline achieved 46.4% sentiment accuracy on out-of-vocabulary text (Satapathy, 2021). This is evidence of linguistic processing difficulty, not a trademark-matching accuracy estimate. Its relevance lies in the need to test informal spellings and phonetic variation rather than assuming clean text.

Mena Badieh Habib's work combines named-entity extraction and disambiguation for informal text, including uncertainty handling and contextual enrichment across languages and domains (Habib, 2014). For monitoring, such disambiguation offers a methodological route to deciding whether an ambiguous word names the protected business, another organisation or something unrelated. The transfer would require brand-specific validation.

Combining observations across platforms introduces further risk. Shuai Wang's 2025 knowledge-graph dissertation shows how erroneous identity assertions can merge unrelated entities and how transitive relations can propagate errors (Wang, 2025). A proposed trademark case graph should therefore distinguish "same seller", "same logo" and "similar wording"; resemblance alone should not establish common ownership. Overall, contextual identity is an evidence-integration task, not merely another similarity score.

5.2. Sentiment and reputational association

Abdulqader Almars develops hierarchical sentiment analysis that connects aspects, topics and user communities, rather than treating every opinion as a flat positive or negative label (Almars, 2019). This structure is useful for distinguishing criticism of a product feature from concerns about authenticity or an alleged commercial association. The application to dilution remains proposed, but the methodological advantage is clear: preserve the target and reason for the sentiment.

Reputation research establishes why broader context can matter. A 2019 working paper uses the 2015 Volkswagen scandal as a natural experiment, combining difference-in-differences and demand estimation. It reports spillovers to non-VW German manufacturers equivalent to an average US$2,057 decline in consumer vehicle valuations and a 34.6% reduction in annual sales (Bachmann et al., 2019). This was collective reputation damage, not evidence of copied marks. It shows that consequential brand-related harm can arise through associations beyond identifier resemblance, while also demonstrating why such harm should not automatically be classified as trademark dilution.

6. Prediction: moving from a detected event to a future outcome

6.1. Economic effects resist a universal risk score

Yi Qian's 2011 working paper combines a natural policy experiment, randomised laboratory experiments and Chinese footwear panel data from 1993–2004. Counterfeiting produced both advertising and substitution effects: advertising dominated for high-end authentic products, whereas substitution dominated for low-end products (Qian, 2011). These findings do not make counterfeiting acceptable; they challenge the empirical assumption that more detected copies always imply the same commercial damage.

Pacula and colleagues offer a different predictive approach. Their two-stage model uses firm forecasts and actual sales for 45 related products from one firm across 16 countries during 2006–2011, first considering observable forecasting errors and then market susceptibility to infringement. It generally underpredicted counterfeit levels but produced trends broadly consistent with shadow-shopping estimates (Pacula et al., 2012). The contrast is instructive: a model can track direction while misestimating magnitude. Monitoring systems should distinguish risk ranking, event probability and expected loss instead of judging all three by one accuracy measure.

6.2. A proposed architecture for dilution-risk forecasting

Figure 1 integrates the detection and interpretation stages discussed above. Its arrows show proposed information flows, not demonstrated causal relationships or a validated dilution predictor.

The retrieval component draws on multimodal trademark search (Ravela & Manmatha, 2005). The contextual component draws on hierarchical sentiment modelling (Almars, 2019). For automated measurement used in downstream estimation, recent econometric work requires combining language-model outputs with a validation sample (Ludwig et al., 2025). The proposed forecast stage should likewise be developed with an independently labelled validation set.

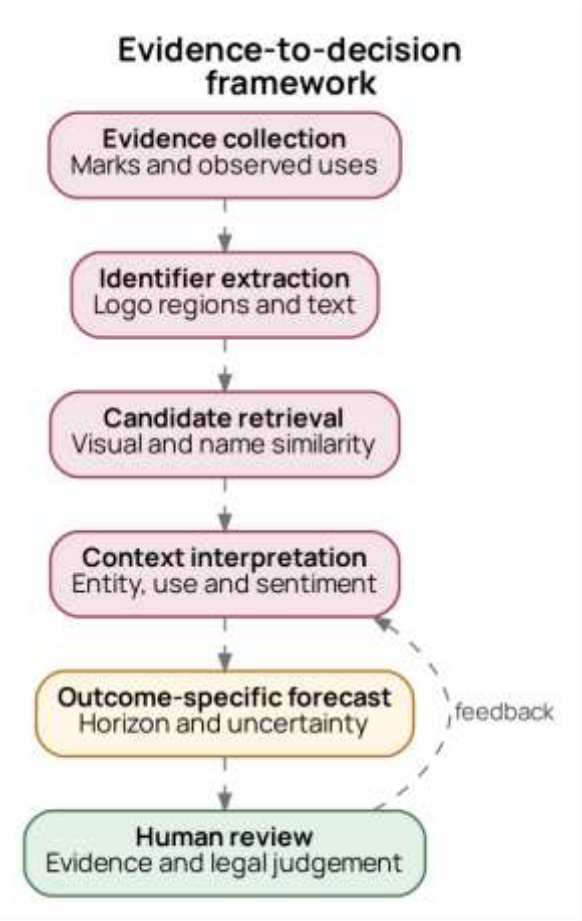


Figure 1. Proposed evidence-to-decision architecture. Dashed arrows denote proposed integration; detecting a similar identifier is not itself a finding of dilution.

A useful forecast must specify its outcome and horizon. Candidate outcomes include expert-adjudicated concern within a defined period, deterioration in a measured consumer association, or a commercial loss measure. These should be modelled separately. As a proposed design, predictors could include identifier resemblance, nature of the use, repeated appearances, observed exposure and changing consumer responses. The system should disclose which evidence is actually observed and which values are proxies. The distinction between measurement and prediction is central. Ludwig, Mullainathan and Rambachan's 2025 working paper argues that text-based prediction requires avoiding training leakage, while downstream estimation using language-model measurements requires a validation sample; model and prompt choices can otherwise change estimates dramatically (Ludwig et al., 2025). Thus, automatically labelling posts “damaging” does not establish that the label measures reputational injury reliably.

6.3. Temporal validity and live monitoring

A 2026 working paper/preprint on point-in-time language models addresses lookahead bias through chronologically restricted training. It reports monthly checkpoints spanning 2013–2024, with models of up to four billion parameters trained on one trillion filtered tokens (Kelly et al., 2026). Its preliminary contribution is methodological: restricting the evaluation dataset by date may be insufficient if the language model already contains later information.

Live monitoring raises another issue. Firas Bayram's 2025 thesis integrates drift detection, real-time data-quality assessment and machine-learning operations within adaptive industrial systems (Bayram,

2025). Applied to trademarks, the proposed requirement is to track whether changing platforms, languages or product presentation alter alert quality. Taken together, these recent contributions make temporal integrity and ongoing validation part of prediction itself, not optional maintenance after deployment.

7. Methodological comparison and evaluation

Table 1 distinguishes directly relevant studies from adjacent evidence. Their datasets and outcomes differ, so neither their sample sizes nor their reported results support a shared performance ranking.

Study	Design or method	Setting or data	Contribution and transfer limit	Citation
Ravela and Manmatha; indexed 2005	Text plus handcrafted visual retrieval	2,000-image evaluation; 63,000-trademark scalability demonstration	Direct trademark search; not a dilution-outcome evaluation	(Ravela & Manmatha, 2005)
Parajuli; 2020	Enhanced autoencoder	Logos within images and transformed variants	Direct logo recognition; improvement reported qualitatively	(Parajuli, 2020)
Song; 2020	Synthetic-data CNN detector	Rendered trademark images; robot demonstration	Trademark-derived training; not enforcement deployment	(Song, 2020)
Sun; 2021	Open-set recognition	Six standard image datasets and train-door pressure data	Unknown-input rejection; trademark transfer untested	(Sun, 2021)
Qian; 2011	Policy natural experiment, laboratory experiments and panel analysis	Chinese footwear; 1993–2004	Heterogeneous counterfeit effects; not legal dilution	(Qian, 2011)
Pacula and colleagues; 2012	Two-stage counterfeit-market model	45 products; 16 countries; 2006–2011	Tracks trends but generally underestimates levels	(Pacula et al., 2012)
Ludwig and colleagues; 2025	Econometric methodological framework	Language-model prediction and measurement	Leakage control and validation requirements; not a trademark trial	(Ludwig et al., 2025)

The review recommends layered evaluation. **Detection** should assess localisation and retrieval, including performance on held-out brands. **Context interpretation** should measure whether annotators and models distinguish relevant commercial uses from unrelated references. **Forecasting** should report

temporal discrimination and probability calibration for a clearly defined outcome. **Operational assessment** should measure analyst workload, review time and the consequences of both missed and unnecessary alerts. These are proposed evaluation criteria, not results already demonstrated by the studies above.

Sebastian Gruber's 2025 dissertation develops uncertainty quantification and calibration through proper scoring rules across classification and other tasks (Gruber, 2025). That provides a methodological basis for examining probability quality rather than relying exclusively on correct class assignments. Alberto Barbado González's anomaly-detection research also evaluates explanations through stability and diversity, incorporating prior domain knowledge in industrial applications (González, 2023). Its implication for trademark review is a testable requirement: explanations should remain consistent under irrelevant changes and identify why the alert warrants attention, without presenting model attribution as a legal justification.

8. Research gaps and integration priorities

The strongest next steps arise from mismatches between the tasks studied, rather than an assertion that the entire field lacks research.

Connect detection to construct-valid outcomes. The transformed-logo research evaluates recognition (Parajuli, 2020), whereas the consumer research distinguishes comprehension, recall and emotional impact (Spencer, 2011). An integrated study should test which detection and contextual features predict a specifically measured change in brand association, keeping legal adjudication separate.

Validate multilingual and cross-platform transfer. Informal-language disambiguation offers transferable methods (Habib, 2014), while out-of-distribution research identifies contextual bias as a generalisation problem (Qi, 2023). Their combination motivates evaluations across scripts, transliterations, platforms and product categories. For an Indian deployment, multilingual coverage would be a proposed validation objective, not a performance claim supported by this review.

Measure intervention value without conflating it with forecast quality. Counterfeit-market models can capture trends while underestimating levels (Pacula et al., 2012). Research should therefore compare prioritisation strategies prospectively and assess whether analyst action improves the chosen outcome. Detection accuracy alone cannot answer that intervention question.

Preserve provenance and reviewability. Knowledge-graph integration can propagate mistaken identity links (Wang, 2025). Case records should retain the source material, timestamps, model version, uncertainty and the analyst's reasoning. This makes disputed alerts revisable rather than treating automated associations as irreversible facts.

9. Conclusion

AI-based trademark monitoring is best understood as a sequence of distinct evidence tasks. Retrieval research supports finding relevant designs and handling transformed logos; language and multimodal methods offer routes to interpreting ambiguous identifiers and surrounding content. Economic research shows that commercial and reputational effects depend on context and cannot be inferred uniformly from the presence of copying. Recent methodological work strengthens the case for validated measurements, temporally credible forecasting and continuous reliability assessment.

The defensible endpoint is therefore not an autonomous verdict of dilution. It is an auditable system that separates resemblance, contextual concern, predicted harm and legal judgement, and supplies analysts

with the evidence needed to examine each. Its effectiveness must be established through outcome-specific validation and prospective assessment of decisions, not assumed from improvements in recognition alone.

References

1. Almars, A. M. A. (2019). *Structured sentiment analysis in social media* [The University of Queensland]. https://doi.org/10.14264/uql.2020.36
2. Ardill-Walker, Janine. (1994). *The economics of trademark infringement* [The University of Queensland]. https://doi.org/10.14264/220066
3. Bachmann, R., Ehrlich, G., Fan, Y., Ruzic, D., & Leard, B. (2019). Firms and collective reputation: A study of the volkswagen emissions scandal. *National Bureau of Economic Research*. https://doi.org/10.3386/w26117
4. Bayram, F. (2025). *Machine learning in motion: Engineering self-adaptive systems*. https://doi.org/10.59217/rjrp8643
5. Bian, Y. (2024). *Towards robust design of deep classification models*. https://doi.org/10.31274/td-20250502-71
6. Chen, Z. (2023). *Generalising to the unseen-zero-shot learning via generative modelling for visual recognition* [The University of Queensland]. https://doi.org/10.14264/9d168dd
7. González, A. B. (2023). *Towards explainable machine learning for anomaly detection in real-world contexts*. https://doi.org/10.20868/upm.thesis.73247
8. Gruber, S. (2025). *A novel framework for uncertainty quantification via proper scores for classification and beyond* [ArXiv.org]. https://doi.org/10.21248/gups.93204
9. Habib, M. B. (2014). *Named entity extraction and disambiguation for informal text: The missing link*. https://doi.org/10.3990/1.9789036536479
10. Kelly, B., Malamud, S., Schwab, J., & Xu, T. A. (2026). Scaling point-in-time language models. *National Bureau of Economic Research*. https://doi.org/10.3386/w35247
11. Liu, L. (2021). *Towards efficient similarity search with semantic hashing techniques* [The University of Queensland]. https://doi.org/10.14264/4c092e0
12. Liu, X. (2016). *Moving up or down : Power distance belief and the asymmetric effects of vertical brand extension*. https://doi.org/10.32657/10356/67027
13. Liu, Z. (2020). *Deep learning methods for scene text detection*. https://doi.org/10.32657/10356/142533
14. Ludwig, J., Mullainathan, S., & Rambachan, A. (2025). Large language models: An applied econometric framework. *National Bureau of Economic Research*. https://doi.org/10.3386/w33344
15. Pacula, R. L., Kadiyala, S., Hunt, P., & Malchiodi, A. (2012). An alternative framework for empirically measuring the size of counterfeit markets. *National Bureau of Economic Research*. https://doi.org/10.3386/w18171
16. Parajuli, E. (2020). *Content based image retrieval (CBIR) for brand logos*. https://doi.org/10.18122/td/1665/boisestate
17. Qi, J. (2023). *Removing bias for out-of-distribution generalization*. https://doi.org/10.32657/10356/168654
18. Qian, Y. (2011). Counterfeiters: Foes or friends? How do counterfeits affect different product quality tiers? *National Bureau of Economic Research*. https://doi.org/10.3386/w16785

19. Ravela, S., & Manmatha, R. (2005). *Multi-modal retrieval of trademark images using global similarity TITLE2:* https://doi.org/10.21236/ada440302
20. Satapathy, R. (2021). *Gauging online and offline public opinion for social media monitoring*. https://doi.org/10.32657/10356/151548
21. Song, J. (2020). *A perception pipeline exploiting trademark databases for service robots* [The University of Queensland]. https://doi.org/10.14264/uql.2020.171
22. Spencer, K. L. (2011). *Evaluating trademark design*. https://doi.org/10.31979/etd.mq3z-gynk
23. Sun, X. (2021). *A study on open set recognition methods*. https://doi.org/10.32657/10356/152461
24. Wang, S. (2025). *Links in Large Integrated Knowledge Graphs: Analysis, Refinement, and Domain Applications. (Liens dans les grands graphes de connaissances intégrés / Liens dans les grands graphes de connaissances intégrés: Analyse, amélioration et applications par domaine)*. https://doi.org/10.5463/thesis.1504
25. Yue, L. (2016). *The myth of the Chinese well-known mark : Formation, debunking and judicial practice*. https://doi.org/10.5353/th_b5774072
26. 陸大瑜. (2010). *Analysis of the assessment of a likelihood of confusion for the community trade mark*. https://doi.org/10.6842/nctu.2010.00456