Novelty Assessment Report

Paper: Difficult Examples Hurt Unsupervised Contrastive Learning: A Theoretical Perspective
PDF URL: https://openreview.net/pdf?id=5LMdnUdAoy
Venue: ICLR 2026 Conference Submission
Year: 2026
Report Generated: 2026-01-07

Abstract

Unsupervised contrastive learning has shown significant performance improvements in recent years, often approaching or even rivaling supervised learning in various tasks. However, its learning mechanism is fundamentally different from supervised learning. Previous works have shown that difficult examples (well-recognized in supervised learning as examples around the decision boundary), which are essential in supervised learning, contribute minimally in unsupervised settings. In this paper, perhaps surprisingly, we find that the direct removal of difficult examples, although reduces the sample size, can boost the downstream classification performance of contrastive learning. To uncover the reasons behind this, we develop a theoretical framework modeling the similarity between different pairs of samples. Guided by this framework, we conduct a thorough theoretical analysis revealing that the presence of difficult examples negatively affects the generalization of contrastive learning. Furthermore, we demonstrate that the removal of these examples, and techniques such as margin tuning and temperature scaling can enhance its generalization bounds, thereby improving performance. Empirically, we propose a simple and efficient mechanism for selecting difficult examples and validate the effectiveness of the aforementioned methods, which substantiates the reliability of our proposed theoretical framework.

Core Task Landscape

This paper addresses: **impact of difficult examples on unsupervised contrastive learning generalization**

A total of **50 papers** were analyzed and organized into a taxonomy with **17 categories**.

Taxonomy Overview

The research landscape has been organized into the following main categories:

- **Hard Negative Sampling Strategies**
- **Supervised and Semi-Supervised Contrastive Learning with Hard Examples**
- **Theoretical Analysis and Mechanisms of Difficulty in Contrastive Learning**
- **Applications and Task-Specific Adaptations**
- **Advanced Contrastive Learning Frameworks and Architectures**

Complete Taxonomy Tree

- impact of difficult examples on unsupervised contrastive learning generalization Survey Taxonomy
- Hard Negative Sampling Strategies
  - Similarity-Based Hard Negative Selection (5 papers)
  - [1] Hard negative sampling strategies for contrastive representation learning (Tabassum, 2022) View paper
  - [21] Hard-Negatives Focused Self-Supervised Learning (Yiming Pan, 2023) View paper
  - [24] Hard Negative Sampling for Contrastive Representation Learning (Jiang, 2024) View paper
  - [30] Contrastive learning with hard negative samples (Robinson, 2020) View paper
  - [40] Conditional Negative Sampling for Contrastive Learning of Visual Representations (Mike Wu, 2021) View paper
  - Generative and Synthetic Hard Negative Creation (4 papers)
  - [2] Hard Negative Mixing for Contrastive Learning (Yannis Kalantidis, 2022) View paper
  - [18] Synthetic Hard Negative Samples for Contrastive Learning (Hengkui Dong, 2024) View paper
  - [27] Contrastive learning with adversarial examples (Chih-Hui Ho, 2020) View paper
  - [48] Improving Knowledge Graph Completion with Generative Hard Negative Mining (Zile Qiao, 2023) View paper
  - Uncertainty and Representativeness-Aware Sampling (3 papers)
  - [10] Difficulty-Based Sampling for Debiased Contrastive Representation Learning (Tae-Uk Jang, 2023) View paper
  - [38] Momentum Contrastive Learning with Enhanced Negative Sampling and Hard Negative Filtering (Hoàng Duy, 2025) View paper
  - [44] Affinity Uncertainty-Based Hard Negative Mining in Graph Contrastive Learning (Chaoxi Niu, 2024) View paper
  - Domain-Specific Hard Negative Mining (8 papers)
  - [6] HAVANA: Hard negative sample-aware self-supervised contrastive learning for airborne laser scanning point cloud semantic segmentation (Yunsheng Zhang, 2024) View paper
  - [16] Effective Hard Negative Mining for Contrastive Learning-Based Code Search (Ye Fan, 2025) View paper
  - [23] NV-Retriever: Improving text embedding models with effective hard-negative mining (Moreira, 2024) View paper
  - [31] ProGCL: Rethinking Hard Negative Mining in Graph Contrastive Learning (Xia Jun, 2021) View paper
  - [33] Contrastive representation learning for self-supervised taxonomy completion (Rui Wang, 2024) View paper
  - [39] HNCSE: Advancing Sentence Embeddings via Hybrid Contrastive Learning with Hard Negatives (Liu WenXiao, 2024) View paper
  - [42] Towards Expansive and Adaptive Hard Negative Mining: Graph Contrastive Learning via Subspace Preserving (Zhezheng Hao, 2024) View paper

- ◦ [47] Self-contrastive learning with hard negative sampling for self-supervised point cloud learning (Du, 2021) View paper
- • Supervised and Semi-Supervised Contrastive Learning with Hard Examples
  - ◦ Supervised Contrastive Learning with Hard Negatives (3 papers)
  - ◦ [3] When hard negative sampling meets supervised contrastive learning (Long, 2023) View paper
  - ◦ [15] Supervised contrastive learning with hard negative samples (Ruijie Jiang, 2024) View paper
  - ◦ [29] Transfer easy to hard: Adversarial contrastive feature learning for unsupervised person re-identification (Haoxuanye Ji, 2024) View paper
  - ◦ Class Imbalance and Long-Tailed Recognition (2 papers)
  - ◦ [5] Parametric contrastive learning (Cui, 2021) View paper
  - ◦ [20] Generalized Parametric Contrastive Learning (Jiequan Cui, 2023) View paper
  - ◦ Semi-Supervised and Positive-Unlabeled Learning (3 papers)
  - ◦ [17] DyCON: Dynamic Uncertainty-aware Consistency and Contrastive Learning for Semi-supervised Medical Image Segmentation (Naseer, 2025) View paper
  - ◦ [35] Teacher-Student Collaboration: Effective Semi-Supervised Model for Defect Instance Segmentation (Biaohua Ye, 2024) View paper
  - ◦ [36] Weighted Contrastive Learning With Hard Negative Mining for Positive and Unlabeled Learning (Botai Yuan, 2025) View paper
- • Theoretical Analysis and Mechanisms of Difficulty in Contrastive Learning
  - ◦ Impact of Difficult Examples on Generalization ★ (2 papers)
  - ◦ [0] Difficult Examples Hurt Unsupervised Contrastive Learning: A Theoretical Perspective (Anon et al., 2026) View paper
  - ◦ [22] Can contrastive learning avoid shortcut solutions? (Robinson, 2021) View paper
  - ◦ Contrastive Loss Behavior and Temperature Effects (2 papers)
  - ◦ [4] Solving inefficiency of self-supervised representation learning (Wang, 2021) View paper
  - ◦ [12] Understanding the behaviour of contrastive loss (Feng Wang, 2021) View paper
  - ◦ Neural Collapse and Representation Geometry (2 papers)
  - ◦ [28] Hard-Negative Sampling for Contrastive Learning: Optimal Representation Geometry and Neural-vs Dimensional-Collapse (Jiang, 2024) View paper
  - ◦ [32] On neural and dimensional collapse in supervised and unsupervised contrastive learning with hard negative sampling (Jiang, 2023) View paper
  - ◦ False Negatives and Sampling Bias (2 papers)
  - ◦ [7] Bayesian self-supervised contrastive learning (Liu Bin, 2023) View paper
  - ◦ [9] Understanding negative samples in instance discriminative self-supervised representation learning (Kento Nozawa, 2021) View paper
- • Applications and Task-Specific Adaptations
  - ◦ Clustering and Unsupervised Categorization (1 papers)
  - ◦ [14] CLC: Cluster Assignment via Contrastive Representation Learning (Ding Fei, 2023) View paper
  - ◦ Domain Adaptation and Transfer Learning (2 papers)
  - ◦ [37] Divide and contrast: Self-supervised learning from uncurated data (Tian, 2021) View paper
  - ◦ [43] Domain confused contrastive learning for unsupervised domain adaptation (Long, 2022) View paper
  - ◦ Specialized Task Applications (5 papers)
  - ◦ [8] Hard negative sample mining for contrastive representation in reinforcement learning (Qihang Chen, 2022) View paper
  - ◦ [11] Improving vision-language models through intra-modal contrastive learning-based hard sample mining (Chaojun Lin, 2025) View paper
  - ◦ [25] Active Learning by Acquiring Contrastive Examples (Barrault, 2021) View paper
  - ◦ [26] Learn from Failure: Causality-guided Contrastive Learning for Generalizable Implicit Hate Speech Detection (Jiang, 2025) View paper
  - ◦ [41] SPContrastNet: A self-paced contrastive learning model for few-shot text classification (Junfan Chen, 2024) View paper
- • Advanced Contrastive Learning Frameworks and Architectures
  - ◦ Hierarchical and Multi-Level Contrastive Learning (3 papers)
  - ◦ [19] Select your own counterparts: Self-supervised graph contrastive learning with positive sampling (Ze-hong Wang, 2024) View paper
  - ◦ [45] Hierarchically contrastive hard sample mining for graph self-supervised pretraining (Wenxuan Tu, 2023) View paper
  - ◦ [46] Self-supervised Graph-level Representation Learning with Adversarial Contrastive Learning (Xiao Luo, 2023) View paper
  - ◦ Self-Contrastive and Intra-Sample Learning (2 papers)
  - ◦ [34] HNSSL: Hard Negative-Based Self-Supervised Learning (Wentao Zhu, 2023) View paper
  - ◦ [50] Revisiting Hard Negative Mining in Contrastive Learning for Visual Understanding (Hao Zhang, 2023) View paper
  - ◦ Universal and Time-Series Representation Learning (2 papers)
  - ◦ [13] Timesurl: Self-supervised contrastive learning for universal time series representation learning (Jiexi Liu, 2024) View paper
  - ◦ [49] Energy-based hopfield boosting for out-of-distribution detection (Sepp Hochreiter, 2024) View paper

Narrative

Core task: impact of difficult examples on unsupervised contrastive learning generalization. The field has organized itself around several complementary perspectives on how example difficulty shapes representation quality. Hard Negative Sampling Strategies form a dense branch exploring algorithmic approaches to identify and prioritize informative negatives, with works like Hard Negative Sampling[1] and Hard Negative Mixing[2] proposing mining techniques that balance difficulty and diversity. Supervised and Semi-Supervised Contrastive Learning with Hard Examples extends these ideas to settings where label information guides the selection process, as seen in Hard Negatives Supervised[3] and HNSSL[34]. Theoretical Analysis and Mechanisms of Difficulty in Contrastive Learning investigates the underlying principles governing when and why difficult examples help or harm, including studies like Understanding Negative Samples[9] and Understanding Contrastive Loss[12]. Applications and Task-Specific Adaptations demonstrate how difficulty-aware strategies translate to domains such as retrieval, graph learning, and vision, while Advanced Contrastive Learning Frameworks and Architectures propose novel model designs that inherently manage example difficulty through architectural choices or dynamic weighting schemes.

A central tension emerges between works advocating aggressive hard negative mining to accelerate convergence and those cautioning against over-reliance on difficult examples that may introduce noise or shortcuts. For instance, Solving Inefficiency[4] and Parametric Contrastive[5] emphasize efficiency gains from targeted sampling, whereas Avoiding Shortcut Solutions[22] warns of pitfalls when models latch onto spurious correlations in hard cases. Difficult Examples Hurt[0] sits squarely within the theoretical branch examining

generalization impacts, closely aligned with Avoiding Shortcut Solutions[22] in questioning the unconditional benefits of difficulty. Unlike purely algorithmic approaches that assume harder is always better, Difficult Examples Hurt[0] provides a nuanced analysis of when difficult examples degrade downstream performance, complementing empirical sampling strategies with principled insights into the difficulty-generalization trade-off that many applied works navigate implicitly.

Related Works in Same Category

The following **1 sibling paper** share the same taxonomy leaf node with the original paper:

1. Can contrastive learning avoid shortcut solutions?

Authors: Robinson, Joshua, Joshua Robinson, Sun Li, Li Sun, et al. (19 authors total) | **Year/Venue**: 2021 | **URL**: View paper

Abstract

The generalization of representations learned via contrastive learning depends crucially on what features of the data are extracted. However, we observe that the contrastive loss does not always sufficiently guide which features are extracted, a behavior that can negatively impact the performance on downstream tasks via "shortcuts", i.e., by inadvertently suppressing important predictive features. We find that feature extraction is influenced by the difficulty of the so-called instance discrimin...

Relationship Analysis

Both papers belong to the same taxonomy category examining how difficult examples impact unsupervised contrastive learning generalization, sharing a focus on theoretical and empirical analysis of this phenomenon. The original paper provides a theoretical framework using similarity graphs to prove that difficult examples hurt generalization bounds and proposes removal, margin tuning, and temperature scaling as solutions. The candidate paper approaches the problem from a different angle, analyzing feature suppression as a consequence of the InfoNCE loss optimization and proposing implicit feature modification (IFM) to prevent encoders from using shortcut solutions that suppress important features, rather than directly removing difficult examples.

Contributions Analysis

Overall novelty summary. The paper contributes a theoretical framework modeling sample-pair similarity and proves that difficult examples negatively affect generalization in unsupervised contrastive learning, proposing removal and mitigation techniques. It resides in the 'Impact of Difficult Examples on Generalization' leaf under 'Theoretical Analysis and Mechanisms of Difficulty in Contrastive Learning', sharing this leaf with only one sibling paper. This positions the work in a relatively sparse research direction focused specifically on generalization bounds rather than algorithmic sampling strategies, distinguishing it from the denser 'Hard Negative Sampling Strategies' branch containing multiple leaves and over fifteen papers.

The taxonomy reveals neighboring leaves examining 'Contrastive Loss Behavior and Temperature Effects', 'Neural Collapse and Representation Geometry', and 'False Negatives and Sampling Bias', all within the theoretical analysis branch. These adjacent directions explore complementary mechanisms—loss properties, optimal geometry, and sampling artifacts—but do not directly address generalization bounds under difficult example removal. The broader 'Hard Negative Sampling Strategies' branch (four leaves, twenty papers) focuses on algorithmic mining techniques, while 'Supervised and Semi-Supervised Contrastive Learning' (three leaves, eight papers) incorporates label information. The paper's theoretical stance on difficulty-induced generalization harm diverges from these predominantly method-oriented neighbors.

Among twenty-three candidates examined, the first contribution (similarity modeling framework) showed no refutable overlap across ten candidates, and the second contribution (proving difficult examples hurt generalization) likewise found no refutations among ten candidates. The third contribution (mitigation techniques improving bounds) examined three candidates and identified two potentially refutable papers, suggesting prior theoretical work on temperature scaling or margin tuning exists. The limited search scope—top-K semantic matches plus citation expansion—means these statistics reflect a focused sample rather than exhaustive coverage, particularly for the mitigation techniques where overlap appears more substantial.

Based on the twenty-three candidates examined, the core theoretical claims about difficult examples harming generalization appear relatively novel within this search scope, while the mitigation techniques connect to existing work on temperature and margin adjustments. The sparse leaf occupancy and absence of refutations for the primary contributions suggest the generalization-focused theoretical angle is less explored than algorithmic sampling methods. However, the limited search scale and the two refutable candidates for mitigation techniques indicate caution is warranted regarding claims of complete novelty, especially for the proposed remedies.

This paper presents **3 main contributions**, each analyzed against relevant prior work:

Contribution 1: Theoretical framework modeling similarity between sample pairs

Description: The authors introduce a similarity graph framework that characterizes relationships between sample pairs in contrastive learning, specifically distinguishing difficult pairs (containing samples near decision boundaries with higher cross-class similarity) from easy pairs. This framework enables formal analysis of how difficult examples affect generalization.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Self-supervised Heterogeneous Graph Neural Network with Co-contrastive Learning

URL: View paper

Brief Assessment

Heterogeneous Co-Contrastive[59] focuses on heterogeneous graph neural networks with cross-view contrastive learning for graph-structured data, not on modeling similarity graphs between sample pairs in general contrastive learning frameworks.

2. Similarity Preserving Adversarial Graph Contrastive Learning

URL: View paper

Brief Assessment

Similarity Preserving Adversarial[66] focuses on graph contrastive learning with similarity-preserving views for adversarial robustness, not on characterizing difficult vs. easy pairs in general contrastive learning frameworks.

3. Adaptive graph contrastive learning for recommendation

URL: View paper

Brief Assessment

Adaptive Graph Contrastive[65] focuses on graph-based recommendation systems using adaptive contrastive view generators, not on theoretical frameworks for modeling similarity between sample pairs in contrastive learning or analyzing difficult examples near decision boundaries.

4. Are graph augmentations necessary? simple graph contrastive learning for recommendation

URL: View paper

Brief Assessment

Simple Graph Contrastive[64] focuses on graph-based recommendation systems and contrastive learning for user-item representations, not on theoretical frameworks for modeling similarity between sample pairs in general contrastive learning or analyzing difficult examples near decision boundaries.

5. Adversarial graph augmentation to improve graph contrastive learning
URL: View paper

Brief Assessment

Adversarial Graph Augmentation[58] focuses on graph contrastive learning with adversarial augmentation strategies for graph-structured data, not on modeling similarity graphs between sample pairs in general contrastive learning frameworks as described in the original contribution.

6. Deep Graph Contrastive Representation Learning
URL: View paper

Brief Assessment

Deep Graph Contrastive[61] focuses on graph-structured data and contrastive learning for node representations, not on modeling similarity graphs for sample pairs in general contrastive learning or analyzing difficult examples near decision boundaries.

7. PE-GCL: Advancing pesticide ecotoxicity prediction with graph contrastive learning
URL: View paper

Brief Assessment

PE-GCL[62] focuses on pesticide ecotoxicity prediction using graph contrastive learning for molecular compounds, not on theoretical frameworks for modeling sample pair similarities in contrastive learning.

8. A noise-resistant graph neural network by semi-supervised contrastive learning
URL: View paper

Brief Assessment

Noise-Resistant Graph[67] focuses on graph neural networks with noisy labels in semi-supervised learning, not on contrastive learning frameworks or similarity modeling between sample pairs in the context described by the original paper.

9. Hard sample aware network for contrastive deep graph clustering
URL: View paper

Brief Assessment

Hard Sample Clustering[63] focuses on graph clustering with a similarity measure combining attribute and structure embeddings for hard sample mining in contrastive learning, not on analyzing difficult examples' effects on generalization bounds in unsupervised contrastive learning frameworks.

10. An empirical study of graph contrastive learning
URL: View paper

Brief Assessment

Empirical Graph Contrastive[60] focuses on empirical evaluation of graph contrastive learning components (augmentations, contrasting modes, objectives) rather than theoretical frameworks for modeling similarity between sample pairs in general contrastive learning. The candidate does not address difficult examples or similarity graphs characterizing cross-class relationships.

Contribution 2: Theoretical analysis proving difficult examples hurt generalization

Description: The authors derive linear probing error bounds for contrastive learning with and without difficult examples, formally proving that difficult examples lead to worse generalization bounds. They show the error bound increases with the presence of difficult samples and worsens as these samples become more challenging.

This contribution was assessed against **10 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Prompted contrast with masked motion modeling: Towards versatile 3d action representation learning
URL: View paper

Brief Assessment

Prompted Contrast Motion[53] focuses on skeleton-based action recognition using contrastive learning and masked motion modeling. It does not address theoretical analysis of difficult examples or generalization bounds in contrastive learning.

2. Dual contrastive learning for general face forgery detection
URL: View paper

Brief Assessment

Dual Contrastive Forgery[56] focuses on face forgery detection using contrastive learning for image classification, not on theoretical analysis of difficult examples affecting generalization bounds in contrastive learning frameworks.

3. When hard negative sampling meets supervised contrastive learning
URL: View paper

Brief Assessment

Hard Negatives Supervised[3] focuses on hard negative sampling in supervised contrastive learning for fine-tuning, not on theoretical generalization bounds for difficult examples in unsupervised contrastive learning.

4. ProGCL: Rethinking Hard Negative Mining in Graph Contrastive Learning
URL: View paper

Brief Assessment

ProGCL[31] focuses on hard negative mining in graph contrastive learning with message passing, not on deriving generalization bounds for difficult examples in general contrastive learning frameworks.

5. When does contrastive visual representation learning work?
URL: View paper

Brief Assessment

When Contrastive Works[52] focuses on empirical conditions for contrastive learning success (data quantity, domain, quality, task granularity) without theoretical generalization bounds or formal analysis of difficult examples.

6. A theory-driven self-labeling refinement method for contrastive representation learning
URL: View paper

Brief Assessment

Theory-Driven Self-Labeling[54] analyzes how inaccurate labels (not specifically difficult examples) affect contrastive learning generalization. The candidate focuses on label corruption and self-labeling refinement, while the original specifically characterizes difficult examples (samples near decision boundaries) and their impact on generalization bounds through a similarity graph framework.

7. Conditional contrastive domain generalization for fault diagnosis
URL: View paper

Brief Assessment

Conditional Contrastive Domain[57] focuses on domain generalization for fault diagnosis in industrial machinery, addressing distribution shifts across working environments. It does not provide theoretical analysis of difficult examples or generalization bounds in contrastive learning.

8. Generalized Parametric Contrastive Learning
URL: View paper

Brief Assessment

Generalized Parametric[20] focuses on class imbalance and parametric centers for contrastive learning, not on difficult examples near decision boundaries affecting generalization bounds.

9. Exploring balanced feature spaces for representation learning
URL: View paper

Brief Assessment

Balanced Feature Spaces[51] focuses on balanced feature spaces in contrastive learning for imbalanced datasets, not on theoretical bounds for difficult examples affecting generalization in contrastive learning.

10. Hard Negative Mining for Domain-Specific Retrieval in Enterprise Systems
URL: View paper

Brief Assessment

Domain-Specific Retrieval Mining[55] focuses on hard-negative mining for enterprise retrieval systems, not on theoretical generalization bounds for contrastive learning with difficult examples.

Contribution 3: Theoretical demonstration of mitigation techniques improving bounds

Description: The authors theoretically prove that three approaches—directly removing difficult examples, margin tuning, and temperature scaling—can mitigate negative effects of difficult examples by improving generalization bounds through different mechanisms of modifying sample pair similarities.

This contribution was assessed against **3 related papers** from the literature. Papers with potential prior art are analyzed in detail with textual evidence; others receive brief assessments.

1. Adaptive Temperature Distillation method for mining hard samples' knowledge
URL: View paper

Brief Assessment

The candidate paper (Adaptive Temperature Distillation[68]) focuses on knowledge distillation for mining hard samples' knowledge, which is a different problem domain from the original paper's analysis of difficult examples in unsupervised contrastive learning and their impact on generalization bounds.

2. A Unified Theoretical Framework for Understanding Difficult-to-learn Examples in Contrastive Learning
URL: View paper

Prior Art Analysis

Unified Difficult Examples[70] demonstrates that the same three mitigation approaches—removing difficult examples, margin tuning, and temperature scaling—were previously proven to improve generalization bounds through modifying sample pair similarities. The candidate paper presents nearly identical theoretical frameworks, proofs, and mechanisms showing how these techniques mitigate negative effects of difficult examples by improving error bounds, establishing prior work on this contribution.

Evidence

Evidence 1 - **Rationale**: Both papers claim the same theoretical contribution of proving that removing difficult samples, margin tuning, and temperature scaling improve generalization bounds. The text is nearly identical, with only 'difficult samples' vs 'difficult-to-learn samples' differing. - **Original**: we theoretically prove that the existence of difficult samples hurts contrastive learning performances. • we theoretically analyze how possible solutions, i.e. directly removing difficult samples, margin tuning, and temperature scaling, can address the issue of difficult examples by improving the ge... - **Candidate**: we theoretically prove that the existence of difficult-to-learn samples hurts contrastive learning performances. • we theoretically analyze how possible solutions, i.e. directly removing difficult-to-learn samples, margin tuning, and temperature scaling, can address the issue of difficult-to-learn e...

Evidence 2 - **Rationale**: The theoretical claims about how these three techniques improve bounds are stated identically in both papers, demonstrating that Unified Difficult Examples[70] previously established this theoretical contribution. - **Original**: we prove that the most straightforward idea of directly removing difficult examples improves the generalization bounds. further, we also theoretically demonstrate that commonly used techniques such as margin tuning (zhou et al., 2024) and temperature scaling (khaertdinov et al., 2022; kukleva et al.... - **Candidate**: we prove that the most straightforward idea of directly removing difficult-to-learn examples improves the generalization bounds. further, we also theoretically demonstrate that commonly used techniques such as margin tuning (zhou et al., 2024) and temperature scaling (khaertdinov et al., 2022; zhang...

Evidence 3 - **Rationale**: Both papers present the same theoretical result (Corollary 4.1) proving that removing difficult examples improves the error bound, with identical mathematical formulations and explanations. - **Original**: corollary 4.1 provides a theoretical explanation for this counter-intuitive phenomenon based on our established framework. corollary 4.1.denote er as the linear probing error of a contrastive learning model trained on a selected subset removing all difficult examplesdd. then there holds er ≤ 4δ 11-α... - **Candidate**: corollary 4.1 provides a theoretical explanation for this counterintuitive phenomenon based on our established framework.

corollary 4.1. denote er as the linear probing error of a contrastive learning model trained on a selected subset removing all difficult-to-learn examples dd. then there holds er...

Evidence 4 - **Rationale**: The theoretical proof that temperature scaling improves bounds is presented with identical mathematical analysis in both papers, establishing prior work on this specific mechanism. - **Original**: theorem 4.5 shows the linear probing error bound of the temperature scaling loss when trained on data containing difficult examples. specifically, with large n and nd/n→0 , we have et/ew.o. - 1≈o((n d/n)2) and ew.d./ew.o. -1≈o(1/n) . this indicates that, when o(nd)≤o(n 1/2), et/ew.o. ≤e w.d./ew.o., ... - **Candidate**: theorem 4.5 shows the linear probing error bound of the temperature scaling loss when trained on data containing difficult-to-learn examples. specifically, with large n and nd/n → 0, we have et/ew.o. - 1 ≈ o((nd/n)2) and ew.d./ew.o. - 1 ≈ o(1/n). this indicates that, when o(nd) ≤ o(n1/2), et/ew.o. ≤...

---

3. Understanding Difficult-to-learn Examples in Contrastive Learning: A Theoretical Framework for Spectral Contrastive Learning

URL: View paper

Prior Art Analysis

Spectral Contrastive Framework[69] demonstrates that the same three mitigation approaches—removing difficult examples, margin tuning, and temperature scaling—improve generalization bounds through similar mechanisms. The candidate paper provides theoretical proofs showing these techniques mitigate negative effects by modifying sample pair similarities, with specific theorems (4.1, 4.3, 4.5) proving improved error bounds. The theoretical frameworks and mathematical formulations are nearly identical, including the use of similarity graphs, adjacency matrices, and spectral contrastive loss formulations.

Evidence

Evidence 1 - **Rationale**: Both papers claim the same theoretical contribution: proving that removing difficult examples, margin tuning, and temperature scaling improve generalization bounds. The text is nearly identical except for terminology ('difficult samples' vs 'difficult-to-learn samples'). - **Original**: we theoretically prove that the existence of difficult samples hurts contrastive learning performances. • we theoretically analyze how possible solutions, i.e. directly removing difficult samples, margin tuning, and temperature scaling, can address the issue of difficult examples by improving the ge... - **Candidate**: we theoretically prove that the existence of difficult-to-learn samples hurts contrastive learning performances. • we theoretically analyze how possible solutions, i.e. directly removing difficult-to-learn samples, margin tuning, and temperature scaling, can address the issue of difficult-to-learn e...

Evidence 2 - **Rationale**: The theoretical claims about how these three techniques improve bounds are identical in both papers, with the same mechanisms described (modifying similarity between sample pairs from different perspectives). - **Original**: we prove that the most straightforward idea of directly removing difficult examples improves the generalization bounds. further, we also theoretically demonstrate that commonly used techniques such as margin tuning (zhou et al., 2024) and temperature scaling (khaertdinov et al., 2022; kukleva et al.... - **Candidate**: we prove that the most straightforward idea of directly removing difficult-to-learn examples improves the generalization bounds. further, we also theoretically demonstrate that commonly used techniques such as margin tuning [38] and temperature scaling [22; 36; 25] mitigate the negative effects of d...

Evidence 3 - **Rationale**: Both papers present the same theoretical result (Corollary 4.1) with identical mathematical formulation proving that removing difficult examples improves the error bound. - **Original**: corollary 4.1 provides a theoretical explanation for this counter-intuitive phenomenon based on our established framework. corollary 4.1.denote er as the linear probing error of a contrastive learning model trained on a selected subset removing all difficult examplesdd. then there holds er ≤ 4δ 11-α... - **Candidate**: corollary 4.1 provides a theoretical explanation for this counter-intuitive phenomenon based on our established framework. corollary 4.1. denote er as the linear probing error of a contrastive learning model trained on a selected subset removing all difficult-to-learn examples dd. then there holds e...

Evidence 4 - **Rationale**: The theoretical result for margin tuning (Theorem 4.3) is identical in both papers, proving that appropriately chosen margins eliminate the negative impact of difficult examples by matching the error bound without such examples. - **Original**: theorem 4.3 shows that with appropriately chosen margins, the linear probing error bound for the margin tuning loss in the presence of difficult examples becomes equivalent to the standard contrastive loss without such examples, as indicated in theorem 3.3. since equation7>0 , this suggests applying... - **Candidate**: theorem 4.3 shows that with appropriately chosen margins, the linear probing error bound for the margin tuning loss in the presence of difficult-to-learn examples becomes equivalent to the standard contrastive loss without such examples, as indicated in theorem 3.1. since equation 7 > 0, this sugges...

Evidence 5 - **Rationale**: The theoretical result for temperature scaling (Theorem 4.5) is identical in both papers, with the same mathematical formulation proving improved convergence rates and error bounds. - **Original**: theorem 4.5 shows the linear probing error bound of the temperature scaling loss when trained on data containing difficult examples. specifically, with large n and nd/n→0 , we have et/ew.o. - 1≈o((n d/n)2) and ew.d./ew.o. -1≈o(1/n) . this indicates that, when o(nd)≤o(n 1/2), et/ew.o. ≤e w.d./ew.o., ... - **Candidate**: theorem 4.5 shows the linear probing error bound of the temperature scaling loss when trained on data containing difficult-to-learn examples. specifically, with large n and nd/n → 0, we have et/ew.o. - 1 ≈ o((nd/n)2) and ew.d./ew.o. - 1 ≈ o(1/n). this indicates that, when o(nd) ≤ o(n1/2), et/ew.o. ≤...

Appendix: Text Similarity Detection

Textual similarity detection checked 24 papers and found 6 similarity segment(s) across 2 paper(s).

The following **2 paper(s)** were detected to have high textual similarity with the original paper. These may represent different versions of the same work, duplicate submissions, or papers with substantial textual overlap. Readers are advised to verify these relationships independently.

1. A Unified Theoretical Framework for Understanding Difficult-to-learn Examples in Contrastive Learning

Detected in: Contribution: contribution_3

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

---

2. Understanding Difficult-to-learn Examples in Contrastive Learning: A Theoretical Framework for Spectral Contrastive Learning

Detected in: Contribution: contribution_3

⚠ **Note**: This paper shows substantial textual similarity with the original paper. It may be a different version, a duplicate submission, or contain significant overlapping content. Please review carefully to determine the nature of the relationship.

References

- [0] Difficult Examples Hurt Unsupervised Contrastive Learning: A Theoretical Perspective View paper
- [1] Hard negative sampling strategies for contrastive representation learning View paper
- [2] Hard Negative Mixing for Contrastive Learning View paper

- [3] When hard negative sampling meets supervised contrastive learning View paper
- [4] Solving inefficiency of self-supervised representation learning View paper
- [5] Parametric contrastive learning View paper
- [6] HAVANA: Hard negative sample-aware self-supervised contrastive learning for airborne laser scanning point cloud semantic segmentation View paper
- [7] Bayesian self-supervised contrastive learning View paper
- [8] Hard negative sample mining for contrastive representation in reinforcement learning View paper
- [9] Understanding negative samples in instance discriminative self-supervised representation learning View paper
- [10] Difficulty-Based Sampling for Debiased Contrastive Representation Learning View paper
- [11] Improving vision-language models through intra-modal contrastive learning-based hard sample mining View paper
- [12] Understanding the behaviour of contrastive loss View paper
- [13] Timesurl: Self-supervised contrastive learning for universal time series representation learning View paper
- [14] CLC: Cluster Assignment via Contrastive Representation Learning View paper
- [15] Supervised contrastive learning with hard negative samples View paper
- [16] Effective Hard Negative Mining for Contrastive Learning-Based Code Search View paper
- [17] DyCON: Dynamic Uncertainty-aware Consistency and Contrastive Learning for Semi-supervised Medical Image Segmentation View paper
- [18] Synthetic Hard Negative Samples for Contrastive Learning View paper
- [19] Select your own counterparts: Self-supervised graph contrastive learning with positive sampling View paper
- [20] Generalized Parametric Contrastive Learning View paper
- [21] Hard-Negatives Focused Self-Supervised Learning View paper
- [22] Can contrastive learning avoid shortcut solutions? View paper
- [23] NV-Retriever: Improving text embedding models with effective hard-negative mining View paper
- [24] Hard Negative Sampling for Contrastive Representation Learning View paper
- [25] Active Learning by Acquiring Contrastive Examples View paper
- [26] Learn from Failure: Causality-guided Contrastive Learning for Generalizable Implicit Hate Speech Detection View paper
- [27] Contrastive learning with adversarial examples View paper
- [28] Hard-Negative Sampling for Contrastive Learning: Optimal Representation Geometry and Neural-vs Dimensional-Collapse View paper
- [29] Transfer easy to hard: Adversarial contrastive feature learning for unsupervised person re-identification View paper
- [30] Contrastive learning with hard negative samples View paper
- [31] ProGCL: Rethinking Hard Negative Mining in Graph Contrastive Learning View paper
- [32] On neural and dimensional collapse in supervised and unsupervised contrastive learning with hard negative sampling View paper
- [33] Contrastive representation learning for self-supervised taxonomy completion View paper
- [34] HNSSL: Hard Negative-Based Self-Supervised Learning View paper
- [35] Teacher-Student Collaboration: Effective Semi-Supervised Model for Defect Instance Segmentation View paper
- [36] Weighted Contrastive Learning With Hard Negative Mining for Positive and Unlabeled Learning View paper
- [37] Divide and contrast: Self-supervised learning from uncurated data View paper
- [38] Momentum Contrastive Learning with Enhanced Negative Sampling and Hard Negative Filtering View paper
- [39] HNCSE: Advancing Sentence Embeddings via Hybrid Contrastive Learning with Hard Negatives View paper
- [40] Conditional Negative Sampling for Contrastive Learning of Visual Representations View paper
- [41] SPContrastNet: A self-paced contrastive learning model for few-shot text classification View paper
- [42] Towards Expansive and Adaptive Hard Negative Mining: Graph Contrastive Learning via Subspace Preserving View paper
- [43] Domain confused contrastive learning for unsupervised domain adaptation View paper
- [44] Affinity Uncertainty-Based Hard Negative Mining in Graph Contrastive Learning View paper
- [45] Hierarchically contrastive hard sample mining for graph self-supervised pretraining View paper
- [46] Self-supervised Graph-level Representation Learning with Adversarial Contrastive Learning View paper
- [47] Self-contrastive learning with hard negative sampling for self-supervised point cloud learning View paper
- [48] Improving Knowledge Graph Completion with Generative Hard Negative Mining View paper
- [49] Energy-based hopfield boosting for out-of-distribution detection View paper
- [50] Revisiting Hard Negative Mining in Contrastive Learning for Visual Understanding View paper
- [51] Exploring balanced feature spaces for representation learning View paper
- [52] When does contrastive visual representation learning work? View paper
- [53] Prompted contrast with masked motion modeling: Towards versatile 3d action representation learning View paper
- [54] A theory-driven self-labeling refinement method for contrastive representation learning View paper
- [55] Hard Negative Mining for Domain-Specific Retrieval in Enterprise Systems View paper
- [56] Dual contrastive learning for general face forgery detection View paper
- [57] Conditional contrastive domain generalization for fault diagnosis View paper
- [58] Adversarial graph augmentation to improve graph contrastive learning View paper
- [59] Self-supervised Heterogeneous Graph Neural Network with Co-contrastive Learning View paper
- [60] An empirical study of graph contrastive learning View paper
- [61] Deep Graph Contrastive Representation Learning View paper
- [62] PE-GCL: Advancing pesticide ecotoxicity prediction with graph contrastive learning View paper
- [63] Hard sample aware network for contrastive deep graph clustering View paper
- [64] Are graph augmentations necessary? simple graph contrastive learning for recommendation View paper
- [65] Adaptive graph contrastive learning for recommendation View paper
- [66] Similarity Preserving Adversarial Graph Contrastive Learning View paper
- [67] A noise-resistant graph neural network by semi-supervised contrastive learning View paper
- [68] Adaptive Temperature Distillation method for mining hard samples' knowledge View paper
- [69] Understanding Difficult-to-learn Examples in Contrastive Learning: A Theoretical Framework for Spectral Contrastive Learning View paper
- [70] A Unified Theoretical Framework for Understanding Difficult-to-learn Examples in Contrastive Learning View paper