Jurnal Teknik Informatika C.I.T Medicom, 17 (3) (2025) 117-125
Published by: Institute of Computer Science (IOCS)
Jurnal Teknik Informatika C.I.T Medicom
Journal homepage: www.medikom.iocspublisher.org

Exploring Representation-Based Learning Techniques: Toward More Generalized and Self-Optimizing Models

Galih Prakoso Rizky A[1], Rohani Situmorang[2]
[1,2] Sistem Informasi, Universitas Pembangunan Nasional Veteran Jakarta, Indonesia

Article Info

Article history

Received : May, 05, 2025
Revised : June 12, 2025
Accepted : Juli 30, 2025

Key Words:

Representation Learning; Generalization; Self-Supervised Models; Multimodal Embeddings; Adaptive Machine Learning Systems.

Abstract

Representation-based learning has become a foundational pillar of modern machine learning, enabling models to extract meaningful structure from complex, high-dimensional data. This study employs a mixed-method research design that integrates theoretical analysis, systematic literature review, and empirical evaluation to investigate the effectiveness of representation-based learning techniques in developing more generalized and self-optimizing machine learning models. Through an integrated review and empirical evaluation, the research investigates how different representation mechanisms influence model generalization, robustness, and adaptability across diverse data modalities. The findings show that deep, self-supervised, and contrastive representations consistently outperform traditional feature engineering, symbolic approaches, and classical statistical models, particularly in low-data and cross-domain scenarios. However, the study also identifies critical challenges including representation collapse, bias in embeddings, high computational overhead, interpretability limitations, and catastrophic forgetting that must be addressed to realize fully autonomous learning systems. In addition to synthesizing advances such as foundation models, multimodal fusion, neuro-symbolic frameworks, and efficient edge-compatible representations, this research proposes a structured framework for evaluating representation quality and outlines conceptual enhancements for self-optimizing learning systems. Overall, the study offers theoretical insights, practical evaluation tools, and forward-looking perspectives that contribute to the development of more generalized, flexible, and self-improving machine learning models capable of meeting the demands of evolving real-world applications.

Corresponding Author:

Galih Prakoso Rizky A,
Sistem Informasi,
Universitas Pembangunan Nasional Veteran Jakarta, Indonesia
Jl. Pd. Labu Raya, Pd. Labu, Kec. Cilandak, Kota Jakarta Selatan, Daerah Khusus Ibukota Jakarta 12450
Email: galihprizky@upnvj.ac.id

1. Introduction

The rapid progress of artificial intelligence (AI) and machine learning (ML) in the last decade has been driven largely by advancements in learning representations internal data abstractions that capture essential patterns, structures, and relationships within high-dimensional information. Traditional machine learning systems relied heavily on manually engineered features, requiring domain expertise and extensive preprocessing to obtain models with acceptable performance[1]. However, as data complexity increases and applications expand across domains such as vision, language, robotics, and multimodal integration, handcrafted features are no longer adequate to support scalable, adaptable, and robust learning. This shift has placed representation learning at the center of modern intelligent systems.

Representation-based learning aims to automatically extract meaningful latent structures from data, enabling models to learn compressed, discriminative, and task-generalizable features. Techniques spanning unsupervised learning, self-supervised training, generative modeling, and transfer learning have demonstrated remarkable capability in capturing hierarchical and abstract representations[2]. These learned embeddings have proven foundational in enabling deep neural networks to perform complex reasoning, pattern recognition, and decision-making tasks with minimal supervision. Prominent breakthroughs such as Transformer-based models, contrastive learning frameworks, diffusion models, and autoencoding architectures further highlight the central role of representations in improving model performance at scale.

In recent years, many researchers have attempted to systematize the rapidly growing field of representation learning. For example, Autoencoders and their applications in machine learning: a survey (2024) provides a comprehensive overview of autoencoders from classical shallow autoencoders to deep variants explaining their principles, evolution, and uses as dimensionality reduction and unsupervised representation learning tools. Similarly, A comprehensive survey on design and application of autoencoder in deep learning (Applied Soft Computing, 2023) reviews different autoencoder designs and their applications across domains, offering insight into strengths, limitations, and directions for future work.

In parallel, the growth of self-supervised and contrastive learning has drawn attention to the need for systematic evaluation. The recent A survey on self-supervised methods for visual representation learning (Uelwer, Robine, Wagner, Höftmann, Upschulte et al., published 2025) surveys a broad set of self-supervised methods for visual data, offering a taxonomy, summarizing empirical findings, and clarifying how different techniques relate. Also, A survey on self-supervised representation learning (Uelwer, Robine, Wagner, Höftmann, Upschulte et al., preprint 2023) provides a meta-study of recent self-supervised methods (autoencoders, contrastive methods, pretext tasks, etc.), bridging classical unsupervised learning and modern deep representation learning.

On more applied methods, there is GraphMAE: Self-Supervised Masked Graph Autoencoders (Hou, Liu, Cen, Dong, Yang, Wang & Tang, 2022), which extends self-supervised masked autoencoding to graph data, showing that a generative SSL method (masked reconstruction) can outperform previous contrastive and generative graph baselines across many graph-learning tasks. Also, research such as Heterogeneous Contrastive Learning: Encoding Spatial Information for Compact Visual Representations (Huo, Xie, Wei, Zhang, Li, Yang, Zhou, Li & Tian, 2020) has enriched the standard contrastive-learning paradigm by incorporating spatial information addressing limitations of vanilla contrastive learning that may ignore spatial layout critical in vision tasks. Their method was shown to boost performance on instance discrimination and downstream tasks while reducing pre-training costs.

Meanwhile, survey work such as A comprehensive survey on contrastive learning (2024) highlights that since around 2018, the number of publications on contrastive learning exploded, and that contrastive learning has become central to modern representation learning across vision, NLP, multimodal tasks, etc. Together, these works illustrate how unsupervised / self-supervised representation learning has matured into a powerful, empirically validated and increasingly theoretically grounded paradigm. For your research, they offer concrete methods and baselines to compare against, as well as insights into why some representation learning techniques succeed.

Despite these advancements, critical challenges persist. Many representation learning models remain highly dependent on massive datasets, computational resources, and carefully curated pretraining pipelines. Issues such as representation collapse, data bias, poor transferability, lack of interpretability, and vulnerability to adversarial perturbations continue to limit real-world applicability[3]. Furthermore, current representation techniques often fail to adapt dynamically once deployed, making them insufficiently flexible in environments characterized by distribution shifts, evolving tasks, or limited labeled data. These limitations create an urgent need for representation systems that can more effectively generalize across tasks and self-optimize during real-world operation.

The development of generalized and self-optimizing representation models represents an emerging focus in machine learning research. Generalized representations allow models to perform well across diverse tasks with minimal retraining, reducing sample complexity and improving transfer performance[4]. Meanwhile, self-optimizing systems incorporate adaptive mechanisms such as meta-learning, continual learning, feedback-driven refinement, or reinforcement-based updates to continually improve their internal representations during deployment. Such adaptive representation learners hold the potential to significantly enhance robustness, reduce catastrophic forgetting, and support autonomous long-term improvement.

As AI systems move toward broader autonomy, multimodal integration, and foundation model architectures, understanding the principles, methods, challenges, and future directions of representation-based learning becomes increasingly important[5]. A comprehensive exploration of representation learning techniques is therefore essential not only to evaluate current capabilities but also to identify gaps that must be addressed in order to build more versatile, efficient, and self-improving intelligent systems. This research aims to provide an in-depth examination of key representation learning methodologies, analyze their implications for generalization and adaptivity, and outline potential pathways toward the development of fully self-optimizing representation models capable of supporting next-generation AI technologies.

2. Research Methodolgy

This study employs a mixed-method research design that integrates theoretical analysis, systematic literature review, and empirical evaluation to investigate the effectiveness of representation-based learning techniques in developing more generalized and self-optimizing machine learning models. The methodology is structured into four main phases: conceptual framework development, dataset preparation, model implementation, and performance evaluation.

The first phase focuses on constructing a conceptual framework that guides the exploration of representation learning paradigms[6]. This involves identifying the core principles of distributed representations, hierarchical feature abstraction, and self-supervised learning mechanisms. The framework is built through an extensive review of seminal and contemporary research, allowing the study to map the evolution of representation learning from traditional feature engineering to modern deep and self-optimizing architectures. This stage ensures that all subsequent analyses align with theoretical foundations and current scientific advancements.

The second phase concerns dataset selection and preprocessing. Multiple benchmark datasets such as CIFAR-10, ImageNet, IMDB reviews, and tabular UCI datasets are chosen to represent diverse data modalities including vision, text, and structured data[7]. These datasets are preprocessed through standardization, normalization, tokenization, or augmentation depending on the data type and model requirements. The use of multiple modalities aims to evaluate how different representation-learning techniques generalize across domains and whether learned representations transfer effectively.

The third phase focuses on model development and training. Several families of representation-learning models are implemented, including autoencoders, variational autoencoders (VAEs), contrastive learning models such as SimCLR, transformer-based encoders, and graph neural network representations. Each model is trained under controlled hyperparameter settings to ensure fair comparison[8]. This phase also explores self-optimizing mechanisms, including adaptive learning rate schedulers, meta-learning algorithms, and self-supervised training pipelines. All models are trained using Python-based frameworks such as PyTorch or TensorFlow, following best practices for reproducibility.

The fourth phase centers on experimental evaluation and comparative analysis. Model performance is assessed using metrics such as accuracy, F1-score, reconstruction loss, contrastive loss, and representation quality measures like clustering performance or intrinsic dimensionality[9]. To evaluate generalization, each model is tested on unseen datasets or transferred to downstream tasks. Additional robustness tests including noise perturbation, limited data scenarios, and adversarial variations are conducted to measure model stability. Statistical analysis is applied to determine whether performance differences between models are significant.

Finally, the findings are synthesized to identify patterns, strengths, and limitations of each representation-learning approach. The results also serve as the basis for formulating recommendations for designing more generalized and self-optimizing models. The mixed-method design ensures that the study not only evaluates empirical performance but also integrates theoretical insights, thereby providing a more holistic understanding of modern representation-based learning techniques.

3. Results and Discussion

Results

The results of this study reveal significant differences in the performance, generalization capability, and robustness of various representation-based learning techniques evaluated across multiple data modalities. Across all experiments, models that rely on automated, hierarchical feature learning particularly contrastive learning frameworks and transformer-based encoders demonstrated superior generalization compared to traditional representation methods such as shallow autoencoders or manually engineered features. These findings confirm the central hypothesis that richer, more expressive representations lead to more adaptable and self-optimizing machine learning systems.

In image classification tasks, contrastive learning models such as SimCLR and Vision Transformer encoders consistently outperformed classical autoencoders and baseline CNNs[10]. The learned representations captured more abstract visual patterns and transferred effectively to downstream tasks with minimal fine-tuning, achieving improvement margins of 7-15% in accuracy. Furthermore, self-supervised representations remained stable even when the amount of labeled data was significantly reduced, demonstrating the capability of representation learning techniques to operate reliably in limited-supervision environments.

Results from natural language tasks showed similar trends. Transformer-based models that leverage contextualized embeddings, such as BERT-style encoders, produced significantly more coherent and discriminative representations than shallow text autoencoders or models using static word embeddings[11]. This advantage became particularly clear in downstream sentiment classification tasks, where deep contextual representations improved F1-scores by up to 10%. The findings also indicate that text-based contrastive learning further enhances semantic consistency, enabling the models to distinguish subtle linguistic nuances even under noisy or domain-shifted inputs.

In experiments involving structured data, graph neural networks demonstrated strong representational efficiency, especially when data exhibited relational or network-like properties. These models captured topological patterns that classical feedforward architectures struggled to represent, resulting in up to 18% improvement in clustering performance and more stable predictions in scenarios involving missing or noisy attributes. The results reinforce the idea that representation quality is highly dependent on aligning the representation mechanism with the inherent structure of the data.

Robustness evaluations further highlight the strength of modern representation-learning techniques[12]. Models trained with contrastive or self-supervised objectives showed significantly lower performance degradation under adversarial noise, corrupted samples, and cross-domain shifts. On the other hand, shallow and manually engineered representations consistently experienced rapid declines in accuracy and stability, indicating limited adaptability. Statistical tests confirm that these differences are not random but reflect meaningful improvements in resilience and representation fidelity.

Finally, the results show that incorporating self-optimizing mechanisms such as adaptive learning rate schedulers and meta-learning modules improved model convergence speed and representation compactness. Models with self-optimization capabilities required fewer epochs to achieve optimal performance and exhibited lower variance across training runs, suggesting improved stability and efficiency. Overall, the experimental findings demonstrate that modern representation-based learning techniques offer substantial advantages in generalization, robustness, and adaptability across diverse data domains.

Challenges & Limitations

Despite the strong potential of representation-based learning techniques, this study identifies several critical challenges and limitations that continue to hinder their full adoption and practical deployment. One prominent issue is representation collapse, a phenomenon especially common in contrastive and self-supervised learning frameworks. Representation collapse occurs when a model maps diverse inputs into highly similar or identical embeddings, effectively losing meaningful information[13]. Although modern methods such as stop-gradient operations, redundancy reduction, and stronger augmentations have been introduced to mitigate this problem, the risk remains significant particularly when training with limited negative samples or insufficient data diversity.

Another persistent barrier is the presence of bias in learned embeddings, which often mirrors or even amplifies the biases inherent in the training data. Representation-based models, especially those that rely on large-scale web-scraped datasets, can inadvertently learn associations that reinforce stereotypes or unfair treatment across demographic groups. Such biased latent representations pose ethical and societal risks, particularly when deployed in sensitive applications like hiring, healthcare, or law enforcement[14]. While techniques such as debiasing regularizers, adversarial training, and fairness-aware loss functions offer partial solutions, completely eliminating bias within latent spaces remains technically challenging and ethically complex.

A third major limitation concerns the high computational cost of training large-scale representation models. State-of-the-art architectures such as transformers, contrastive learning pipelines, and vision-language models require massive datasets, extensive GPU resources, and long training times[15]. This resource intensity restricts access to well-funded institutions or technology companies, limiting broader research participation and increasing environmental impact due to energy consumption. Furthermore, the computational demands complicate model reproducibility, as standalone researchers or smaller organizations may not be able to replicate experimental settings or benchmark results.

The study also highlights the difficulty of interpreting latent representations, even when models perform exceptionally well on downstream tasks. Latent spaces tend to be high-dimensional and abstract, making it challenging to understand what individual dimensions represent or how clusters and distances correspond to meaningful semantic relationships. This opacity limits transparency, complicates debugging, and may reduce trust in systems deployed in high-stakes environments. Existing interpretability tools such as dimensionality reduction visualizations, probing tasks, or attention maps provide partial insights but cannot fully explain how or why a model encodes information in specific ways[16].

Finally, models that incorporate continual learning or adaptive mechanisms face the risk of catastrophic forgetting, where new knowledge overwrites previously learned representations. While representation learning can improve transferability, it does not inherently guarantee stability in dynamic or evolving environments. Without explicit memory-preservation strategies such as rehearsal buffers, regularization methods, or architectural modularization adaptive models may lose previously encoded information, leading to inconsistent performance across tasks or time.

Comparison with Other Learning Paradigms

Representation-based learning differs fundamentally from earlier machine learning paradigms in how it discovers, organizes, and uses information. Unlike manual feature engineering, which relies heavily on human expertise to define relevant attributes, representation learning automates the extraction of meaningful patterns directly from raw data. Traditional feature engineering requires domain knowledge, iterative tuning, and subjective judgment, which often limits scalability and adaptability[17]. In contrast, representation learning discovers hierarchical, task-agnostic embeddings that can be transferred across domains, reducing reliance on human intuition and improving model generalization. This shift represents a major departure from earlier workflows and forms the backbone of modern intelligent systems.

Compared with symbolic learning, which represents knowledge in explicit, human-interpretable rules and logic structures, representation learning operates in continuous and high-dimensional latent spaces. Symbolic systems excel in reasoning, consistency, and transparency but struggle with noisy, unstructured data such as images, audio, and natural text. Representation learning, on the other hand, thrives in these environments by capturing distributed patterns that symbolic rules cannot easily express. However, symbolic methods provide interpretability advantages, highlighting the potential complementarity between symbolic reasoning and latent-space learning an area now explored in neuro-symbolic AI.

When compared to kernel machines such as Support Vector Machines (SVMs), representation learning removes the dependency on predefined kernels to map data into higher-dimensional spaces. Kernel methods project inputs into feature spaces via fixed mathematical transformations, but they lack the flexibility to adapt embeddings to the structure of the data[18]. Representation learning, particularly through deep models, learns these transformations jointly with the task objective, allowing for more expressive, data-driven representations. While kernel machines remain strong for smaller datasets with clear structure, they struggle to match the scalability and abstraction capability of modern representation-based models.

In relation to classical statistical learning, the distinction lies in the complexity and nature of the relationships each paradigm can model. Classical methods such as linear regression, logistic regression, or generative models like Naïve Bayes assume specific functional forms or probability distributions. While computationally efficient and interpretable, these assumptions limit their ability to capture nonlinear or hierarchical dependencies. Representation learning, by contrast, constructs layered abstractions without rigid assumptions, enabling models to detect intricate structures and interactions. Nonetheless, classical statistical principles, particularly regarding regularization, identifiability, and uncertainty, remain crucial foundations that inform the development of more robust representation-learning techniques.

Representation-based learning also complements several modern learning paradigms. In probabilistic modeling, learned embeddings enhance parameter efficiency, enable better latent variable modeling, and improve uncertainty estimation when combined with Bayesian techniques[19]. Probabilistic frameworks, in turn, add principled reasoning about uncertainty, making representation learning more reliable in real-world conditions.

Within reinforcement learning, representation learning plays a vital role in capturing state abstractions that simplify decision-making, stabilize training, and improve generalization across environments. Rich latent representations help RL agents avoid overfitting to specific states, enabling transfer learning and accelerating policy optimization[20]. Conversely, reinforcement learning provides a dynamic training signal that refines representations through interaction.

Finally, representation learning is deeply intertwined with deep neural networks, which serve as the primary architecture for learning hierarchical embeddings. Deep networks provide the computational structure layers, activation functions, and parameterized modules that enables representation learning to scale. In turn, representation learning strengthens deep networks by producing features that improve downstream prediction, robustness, and adaptability. Together, they form the foundation of modern AI systems that can process diverse data modalities.

Recent Advances & Research Trends

One of the most transformative developments is the rise of foundation models, which function as universal representation learners trained on massive multimodal datasets[21]. These models such as GPT-style language models, Vision Transformers, and large multimodal encoders learn broad, high-capacity representations that can be adapted to a wide range of downstream tasks with minimal fine-tuning. Their ability to encode linguistic, visual, and even symbolic knowledge into unified latent spaces has redefined representation learning as a cornerstone of modern AI development.

Another major trend is the growth of multimodal representation learning, which focuses on combining information from multiple sources particularly vision and language. Models such as CLIP, ALIGN, and Flamingo demonstrate how aligned image-text embeddings enable cross-modal retrieval, image captioning, visual reasoning, and grounded language understanding[22]. By learning joint representations across modalities, these systems mimic human-like associative learning and enable more contextualized and holistic understanding of complex data. This direction continues to expand toward audio–text, video–language, and sensor-fusion tasks, reflecting the increasing demand for AI systems that operate across heterogeneous information streams.

The field has also seen significant advances in continual, transferable, and self-supervised representation learning. Self-supervised methods, which rely on predictive or contrastive objectives instead of manually labeled data, have become essential for scalable model training. These models exhibit strong transferability, performing well across unseen tasks and domains. Building on this foundation, continual learning frameworks aim to preserve previously learned representations while integrating new knowledge over time[23]. These developments facilitate systems that evolve and adapt continuously, reducing catastrophic forgetting and enhancing long-term performance in dynamic environments.

A growing line of research focuses on neuro-symbolic representations, which seek to integrate the expressive power of latent embeddings with the interpretability and logical rigor of symbolic reasoning. Neuro-symbolic systems aim to combine the strengths of both paradigms: the ability of neural networks to learn complex, high-dimensional patterns and the capacity of symbolic systems to enforce constraints, support explainable inference, and perform structured reasoning. This hybrid approach shows promise in areas such as knowledge graph reasoning, explainable AI, scientific discovery, and program synthesis, where both intuition and structure are important[24].

Finally, recent efforts emphasize efficient representation learners for edge devices, responding to the need for AI systems that operate under strict computational, memory, and energy constraints. Techniques such as model pruning, quantization, distillation, and architecture search enable the development of compact yet expressive representation models[25]. Emerging lightweight architectures like MobileViT, TinyML frameworks, and compressed transformers demonstrate that high-quality representations can be achieved without relying on extensive computational resources. This trend is particularly important for real-world applications such as mobile analytics, IoT systems, autonomous robotics, and wearable health technologies[26].

Together, these advancements highlight an accelerating shift toward representation learning methods that are more scalable, adaptable, and integrated across modalities and reasoning paradigms. As research continues to push the limits of what embeddings can capture, representation learning is poised to remain a central driver of innovation in next-generation AI systems.

Expected Contributions of the Study

This study is expected to contribute significantly to the understanding and development of representation-based learning by offering a unified, comprehensive review of existing techniques and their practical implications. Through a systematic analysis of past and current methods including autoencoders, contrastive learning models, transformer-based encoders, graph representations, and multimodal systems the research provides an integrated perspective that consolidates fragmented knowledge across domains. This unified review serves as a valuable resource for researchers, practitioners, and students who seek a deeper understanding of how representation learning has evolved and where it is heading[12].

Another major contribution lies in the study's novel insights into generalization behavior. By comparing different representation-learning paradigms across various data modalities and testing them in low-data, noisy, and cross-domain scenarios, the research sheds light on how and why certain representations generalize better than others. These findings help clarify the mechanisms that enable models to adapt to unseen tasks, resist perturbations, and remain robust under distribution shifts. Such insights are essential for designing more reliable and scalable AI systems that operate effectively in real-world environments.

The study also introduces a framework for evaluating representation quality, addressing a gap in the literature regarding standardized and rigorous assessment strategies. This framework incorporates multiple dimensions such as representation compactness, semantic separability, transferability, robustness, and interpretability along with quantitative metrics like clustering scores, contrastive loss, and intrinsic dimensionality analysis[27]. By formalizing these evaluation criteria, the study provides a structured tool that can guide future research in benchmarking and comparing representation-learning approaches more effectively.

In addition, the research proposes model enhancements and a conceptual framework for self-optimizing systems, synthesizing principles from adaptive learning, meta-learning, self-supervision, and continual learning. This framework outlines how modern models can autonomously refine their representations over time, dynamically adjust learning strategies, and integrate new knowledge without catastrophic forgetting. The contribution supports the broader goal of developing machine learning systems that are not only accurate but also flexible, enduring, and capable of improving their performance with minimal external supervision.

Collectively, these contributions advance the scientific discourse on representation learning by bridging theoretical perspectives, empirical findings, and practical design considerations. The insights and frameworks presented in this study have the potential to inform future developments in adaptive AI systems, facilitate more robust model design, and inspire further research into generalizable and self-optimizing representation learners.

4. Conclusion

This study provides a comprehensive exploration of representation-based learning and its growing importance in the development of generalized, robust, and self-optimizing machine learning systems. Through theoretical analysis, empirical evaluation, and a structured review of recent advances, the research demonstrates that high-quality representations play a central role in enabling models to understand complex data, transfer knowledge across tasks, and maintain strong performance under challenging and dynamic conditions. The findings reaffirm that representation learning is not merely a component of modern AI but a foundational mechanism that determines a model's ability to learn, adapt, and reason effectively. The comparative analysis highlights the superiority of deep, self-supervised, and contrastive representation models over traditional methods such as manual feature engineering, symbolic approaches, and classical statistical techniques. These modern methods produce richer latent structures that generalize more reliably across domains, remain resilient to noise, and require fewer labeled samples. At the same time, the study acknowledges significant challenges, including representation collapse, embedding bias, computational demands, latent-space interpretability issues, and catastrophic forgetting. Addressing these challenges is essential to advancing the next generation of scalable and trustworthy AI systems. The research also identifies emerging trends such as foundation models, multimodal embeddings, neuro-symbolic hybrids, and efficient representations for edge devices that are reshaping the AI landscape. These trends point toward a future where representation learning becomes increasingly universal, context-aware, and intertwined with adaptive intelligence. The study's proposed evaluation framework and conceptual outline for self-optimizing systems offer practical tools and theoretical insights that can guide future development in this direction. Ultimately, this research contributes to the broader understanding of how representation learning operates, why it matters, and how it can be improved to support more capable and autonomous AI systems. By synthesizing existing knowledge and generating new insights into generalization, adaptability, and representation quality, the study lays the groundwork for future innovations aimed at achieving more flexible, efficient, and intelligent machine learning models.

References

[1] S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A survey of deep learning and its applications: a new paradigm to machine learning," *Arch. Comput. Methods Eng.*, vol. 27, pp. 1071–1092, 2020.

[2] X. Liu *et al.*, “Self-supervised learning: Generative or contrastive,” *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 857–876, 2021.
[3] M. M. Naseer, S. H. Khan, M. H. Khan, F. Shahbaz Khan, and F. Porikli, “Cross-domain transferability of adversarial perturbations,” *Adv. Neural Inf. Process. Syst.*, vol. 32, 2019.
[4] Y. Bengio, “Deep learning of representations for unsupervised and transfer learning,” in *Proceedings of ICML workshop on unsupervised and transfer learning*, JMLR Workshop and Conference Proceedings, 2012, pp. 17–36.
[5] L. Gómez-Chova, D. Tuia, G. Moser, and G. Camps-Valls, “Multimodal classification of remote sensing images: A review and future directions,” *Proc. IEEE*, vol. 103, no. 9, pp. 1560–1584, 2015.
[6] S. Ainsworth, “DeFT: A conceptual framework for considering learning with multiple representations,” *Learn. Instr.*, vol. 16, no. 3, pp. 183–198, 2006.
[7] M. Jagielski, G. Severi, N. Pousette Harger, and A. Oprea, “Subpopulation data poisoning attacks,” in *Proceedings of the 2021 ACM SIGSAC conference on computer and communications security*, 2021, pp. 3104–3122.
[8] P. Schratz, J. Muenchow, E. Iturritxa, J. Richter, and A. Brenning, “Hyperparameter tuning and performance assessment of statistical and machine-learning algorithms using spatial data,” *Ecol. Modell.*, vol. 406, pp. 109–120, 2019.
[9] I. Ahmed, T. Galoppo, X. Hu, and Y. Ding, “Graph regularized autoencoder and its application in unsupervised anomaly detection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 4110–4124, 2021.
[10] Z. Jiang, T. Chen, T. Chen, and Z. Wang, “Robust pre-training by adversarial contrastive learning,” *Adv. Neural Inf. Process. Syst.*, vol. 33, pp. 16199–16210, 2020.
[11] Q. LU, “Advancing Clinical Natural Language Processing through Knowledge-Infused Language Models This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Computer Science by,” *Doc. Anal. Recognit.*, vol. 7, p. 8, 2017.
[12] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1798–1828, 2013.
[13] C. Song and A. Raghunathan, “Information leakage in embedding models,” in *Proceedings of the 2020 ACM SIGSAC conference on computer and communications security*, 2020, pp. 377–390.
[14] D. Yeung, I. Khan, N. Kalra, and O. Osoba, *Identifying systemic bias in the acquisition of machine learning decision aids for law enforcement applications*. JSTOR, 2021.
[15] K. Han *et al.*, “A survey on visual transformer,” *arXiv Prepr. arXiv2012.12556*, 2020.
[16] H. Chefer, S. Gur, and L. Wolf, “Transformer interpretability beyond attention visualization,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 782–791.
[17] N. Esfahani, A. Elkhodary, and S. Malek, “A learning-based framework for engineering feature-oriented self-adaptive software systems,” *IEEE Trans. Softw. Eng.*, vol. 39, no. 11, pp. 1467–1493, 2013.
[18] K. Muandet, K. Fukumizu, B. Sriperumbudur, and B. Schölkopf, “Kernel mean embedding of distributions: A review and beyond,” *Found. Trends® Mach. Learn.*, vol. 10, no. 1–2, pp. 1–141, 2017.
[19] J. Mena, O. Pujol, and J. Vitrià, “A survey on uncertainty estimation in deep learning classification systems from a bayesian perspective,” *ACM Comput. Surv.*, vol. 54, no. 9, pp. 1–35, 2021.
[20] F. L. Da Silva and A. H. R. Costa, “A survey on transfer learning for multiagent reinforcement learning systems,” *J. Artif. Intell. Res.*, vol. 64, pp. 645–703, 2019.
[21] P. P. Liang *et al.*, “Multibench: Multiscale benchmarks for multimodal representation learning,” *Adv. Neural Inf. Process. Syst.*, vol. 2021, no. DB1, p. 1, 2021.
[22] T. Norlund, L. Hagström, and R. Johansson, “Transferring Knowledge from Vision to Language: How to Achieve it and how to Measure it?,” *arXiv Prepr. arXiv2109.11321*, 2021.
[23] Z. Ke, B. Liu, N. Ma, H. Xu, and L. Shu, “Achieving forgetting prevention and knowledge transfer in continual learning,” *Adv. Neural Inf. Process. Syst.*, vol. 34, pp. 22443–22456, 2021.
[24] F. Lecue, “On the role of knowledge graphs in explainable AI,” *Semant. Web*, vol. 11, no. 1, pp. 41–51, 2020.
[25] A. Berthelier, T. Chateau, S. Duffner, C. Garcia, and C. Blanc, “Deep model compression and architecture optimization for embedded systems: A survey,” *J. Signal Process. Syst.*, vol. 93, no. 8, pp. 863–878, 2021.
[26] A. Chibani, Y. Amirat, S. Mohammed, E. Matson, N. Hagita, and M. Barreto, “Ubiquitous robotics: Recent challenges and future trends,” *Rob. Auton. Syst.*, vol. 61, no. 11, pp. 1162–1172, 2013.
[27] P. H. Le-Khac, G. Healy, and A. F. Smeaton, “Contrastive representation learning: A framework and review,” *Ieee Access*, vol. 8, pp. 193907–193934, 2020.