The Role of GANs in Activity Recognition Methods

Lubah Nelson

Dec 3, 2024

Abstract

Activity recognition is a vital field in artificial intelligence with applications spanning healthcare, surveillance, and human-computer interaction. However, challenges such as data scarcity, multimodal complexity, and domain adaptation hinder progress. Generative Adversarial Networks (GANs) have emerged as a promising solution to these challenges. This survey explores the role of GANs in activity recognition, highlighting their contributions in data augmentation, improving robustness, and facilitating multimodal integration. The paper also discusses current challenges and future directions to enhance GAN-based approaches in this domain.

1 Introduction

Activity recognition is a multidisciplinary field that combines computer vision, machine learning, and sensor technologies. It focuses on the automatic identification and classification of actions and behaviors using diverse data sources, including video, audio, and sensor inputs. Its applications span diverse domains, including healthcare, where systems enable fall detection and rehabilitation tracking [29], surveillance for real-time anomaly detection [30], and industrial automation for predictive maintenance [31].

Despite its versatility, activity recognition faces persistent challenges. Data scarcity limits the availability of representative examples for rare activities, particularly in domains such as healthcare [32]. Domain adaptation issues arise due to variations in environmental conditions and sensor data, leading to significant performance drops in real-world deployment [33]. Moreover, multimodal integration, essential for combining video, audio, and sensor data, presents scalability and computational challenges [34]. Traditional machine learning models and even advanced deep learning architectures like CNNs and RNNs struggle to address these complexities comprehensively, particularly in scenarios requiring large labeled datasets or multimodal processing [35].

Generative Adversarial Networks (GANs) have emerged to tackle these challenges. GANs are a type of generative deep learning technique that use artificial neural networks to create synthetic data. Their adversarial framework consists of two competing networks—a generator, which produces synthetic data, and a discriminator, which evaluates the authenticity of the generated data. Through this iterative process, GANs refine their outputs to produce highly realistic data that closely mimics real-world distributions [36].

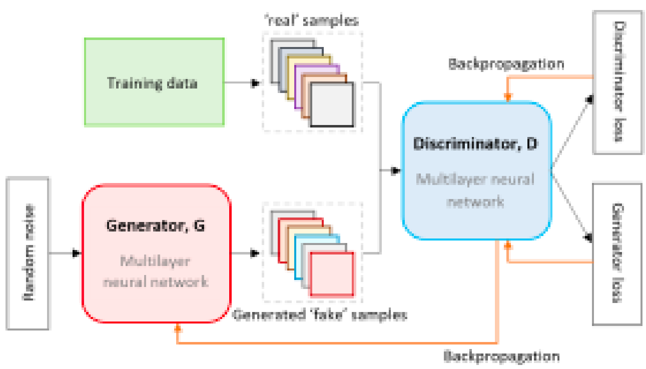

Figure 1: Example of GAN Architecture

Figure 1: An overview of the GAN architecture. Adapted from [8].

By leveraging this framework, GANs excel at addressing the core challenges in activity recognition:

- **Data augmentation**: Mitigating data scarcity by generating diverse synthetic samples for underrepresented classes [45].

- **Domain Adaptation**: Simulating diverse environmental conditions to improve model adaptability to real-world scenarios [12].

- **Multimodal integration**: Synthesizing and aligning disparate data streams to enable the fusion of video, audio, and sensor inputs [25].

This survey explores the role of GANs in activity recognition. Specifically, it addresses these challenges, with a focus on their applications in data augmentation, robustness improvement, and multimodal integration. Additionally, the paper evaluates the limitations of GANs, such as training instability and computational demands, and discusses ongoing advancements that aim to expand their applicability in real-world activity recognition systems.

2 Background

Activity recognition, emerging in the late 1990s, is a multidisciplinary research field aimed at the automatic identification and classification of human and non-human behaviors. Early methods relied heavily on manual feature engineering, leveraging techniques such as motion vectors, optical flow, and template matching to identify patterns in data. While these approaches were effective in controlled environments, their reliance on domain-specific expertise and inability to handle variability in real-world conditions limited their scalability and robustness [1].

The introduction of machine learning (ML) marked a significant advancement, enabling algorithms such as Support Vector Machines (SVMs), Random Forests, and Decision Trees to automatically learn patterns from data. These approaches shifted the paradigm away from manual feature extraction, offering improved accuracy and adaptability in structured environments [2]. Despite their success, ML models remained reliant on handcrafted features and often struggled with scalability and adaptability in diverse, dynamic contexts.

Deep learning revolutionized activity recognition by automating feature extraction and modeling complex patterns through multilayered neural networks. Convolutional Neural Networks (CNNs) are effective for capturing spatial features, making them essential for image-based tasks [3]. However, their focus on spatial hierarchies limits their ability to process temporal dynamics needed for sequential activities like jogging or walking. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks address this limitation by capturing temporal dependencies, enabling accurate recognition of sequential actions [4]. Despite their advancements, RNNs and LSTMs can be inefficient for large-scale or multimodal datasets. Hybrid models, such as CNN-LSTMs, partially resolve these issues by combining spatial and temporal learning. However, they still face challenges with scalability and generalization in real-world conditions, such as noise or sensor misalignment.

The limitations of existing activity recognition approaches become particularly evident when addressing three key challenges as mentioned: data scarcity, domain adaptation, and multimodal complexity. These challenges highlight the barriers that prevent the broader applicability of activity recognition systems in real-world environments.

The first challenge, **data scarcity**, is particularly problematic in domains where rare events or specialized activities are critical. Healthcare datasets, for example, often overwhelmingly represent normal activities, with only a small fraction capturing rare but critical events like falls [11]. This imbalance skews model performance, making it difficult to accurately recognize minority-class events. The difficulty of collecting labeled data for such events further compounds this issue, especially in sensitive domains like healthcare.

A second significant challenge is **domain adaptation**, which involves generalizing models trained in controlled environments to diverse, real-world conditions. Models developed with structured datasets often falter when exposed to variability such as changing lighting, sensor placements, or environmental noise [12]. For instance, a model trained on indoor activities in well-lit environments may fail when deployed outdoors, where lighting and environmental conditions are unpredictable. These discrepancies emphasize the need for approaches capable of bridging the gap between training and deployment contexts.

The third major challenge is **multimodal complexity**, which arises from the integration of diverse data streams. While multimodal systems offer richer contextual information and improve accuracy, they introduce significant technical challenges. Differences in sampling rates and data formats across modalities require sophisticated preprocessing and synchronization, while missing or noisy data from one modality—such as video feeds in low-light conditions—can disrupt overall system performance [13]. However, it is in these situations where GANs outperform these bottlenecks.

Generative Adversarial Networks (GANs) have emerged as a framework for overcoming these bottlenecks. Unlike traditional methods, GANs leverage an adversarial architecture comprising a generator and a discriminator, which collaboratively refine synthetic data to closely mimic real-world distributions. This dynamic interplay enables GANs to produce highly realistic and diverse data, equipping models to handle variability, noise, and data imbalances inherent in real-world applications. By synthesizing representative samples, GANs support activity recognition systems in bridging critical gaps in data availability, enhancing adaptability to environmental complexities, and integrating multimodal inputs. A unique advantage of GANs is their privacy-preserving capability. Unlike traditional data augmentation methods, the generative process in GANs does not require direct access to original data during training. This ensures that sensitive data, such as medical records or personal sensor data, remains protected while enabling the generation of representative synthetic datasets [45].

The versatility of GANs makes them particularly suited for addressing the multifaceted challenges of activity recognition, from rebalancing skewed datasets to improving generalization across diverse contexts. Their ability to generate data reflective of complex patterns and contextual dependencies positions them as a powerful tool for improving robustness and reliability. While GANs are not without their limitations, including issues such as training instability and computational overhead, advancements in architecture design and optimization continue to expand their utility. The following sections delve into the specific mechanisms through which GANs address the persistent challenges of activity recognition—data scarcity, domain adaptation, and multimodal integration—highlighting their transformative potential and practical applications.

2.1 Data Augmentation in Human Activity Recognition

Data augmentation is a fundamental technique in ML, particularly for addressing data scarcity and imbalanced datasets. By generating diverse and realistic synthetic data, augmentation increases the size and variability of training datasets, enabling models to generalize better to unseen scenarios. Traditional augmentation methods—such as geometric transformations, temporal shifts, and noise injection—offer initial solutions but are inherently limited in their ability to create entirely new samples or capture the complexity of real-world distributions [15].

GANs have emerged as a revolutionary approach to data augmentation, overcoming these limitations while maintaining the original data's privacy. GANs generate synthetic data reflective of real-world complexities, effectively addressing the limitations of traditional methods [15, ?]. They are particularly valuable in domains with scarce or sensitive data, as the generative process does not require direct access to the original dataset, preserving privacy [?]. Moreover, the ability to simulate diverse scenarios enables GANs to produce datasets tailored to specific challenges, improving model adaptability [17, 21]. The impact of GANs in data augmentation can be summarized across three critical areas: enhancing generalization, addressing class imbalance, and simulating variability.

2.2 Enhancing Generalization

GANs significantly reduce overfitting by exposing models to a broader range of synthetic examples that capture complex relationships within data, leading to improved performance on unseen scenarios. Unlike traditional augmentation methods, GANs generate data reflective of nuanced patterns and dependencies, enabling models to generalize better. For instance, a GAN-based augmentation approach applied to a fitness dataset improved validation accuracy by 15% [16]. Similarly, in medical imaging, GANs have been instrumental in generating synthetic scans for rare conditions, enhancing diagnostic models' ability to generalize across diverse patient populations [16].

2.3 Addressing Class Imbalance

Class imbalance poses significant challenges in domains where rare events or conditions are critical. GANs excel at generating realistic samples for underrepresented classes, ensuring balanced datasets and improved model performance. For example, in HAR, GANs increased the minority class representation for "falling" from 3% to 30%, leading to a 25% improvement in recall [17]. In fraud detection, GANs have been successfully used to generate synthetic fraudulent transactions, mitigating the imbalance between legitimate and fraudulent cases in training data [18].

2.4 Simulating Variability

Variability in human activities, environmental contexts, and physical conditions often challenges model robustness. GANs address this by introducing controlled variability, enabling models to learn adaptable representations. For instance, GANs modeled individual variations in hand gestures, improving generalization across diverse user populations [19]. Similarly, in autonomous systems, GANs simulate diverse driving scenarios—such as weather changes and varying traffic densities—enhancing navigation models' reliability and performance under real-world conditions [21].

2.5 Enhancing Domain Adaptation

Domain discrepancies, often caused by differences between training and deployment environments, pose significant challenges to model generalization. GANs address these discrepancies by aligning source (training) and target (deployment) domain distributions. This alignment is achieved through adversarial training. Over iterative training, the generator learns to capture the underlying characteristics of the target domain, effectively bridging the gap between the source and target distributions. This process enables models trained on source data to perform effectively in target domains with different data distributions [28].

2.5.1 Environmental Variability

Environmental variability, such as changes in lighting, weather, and sensor placements, often alters data distributions, leading to performance degradation. GANs address this by generating synthetic data that represents diverse environmental conditions.

The generative mechanism of GANs enables the creation of realistic representations of environmental changes by encoding such variability into the latent space. For instance, in activity recognition tasks, GANs have been employed to simulate low-light scenarios, enhancing the model's ability to recognize activities during nighttime. Studies leveraging the MultiTHUMOS dataset demonstrated that GAN-generated nighttime activity data improved accuracy by up to 20

In autonomous vehicle systems, GANs simulate diverse weather scenarios—such as fog, rain, and snow—providing training datasets that capture the nuanced effects of adverse conditions. By incorporating these variations into the training process, GAN-augmented datasets improved detection rates of pedestrians and vehicles in challenging weather environments [21]. These simulations refine latent space embeddings to encapsulate the distributional properties of target environments.

2.5.2 Mitigating Sensor Noise

Sensor noise, arising from hardware limitations, external interference, or individual user behaviors, presents a significant barrier to accurate predictions. GANs address this issue by embedding realistic noise patterns into synthetic data, enabling models to learn features that are robust to such imperfections.

For example, Noise-Conditional GANs (NC-GANs) specialize in this area by generating synthetic datasets conditioned on specific noise profiles. These datasets ensure realistic distortions while preserving underlying activity patterns. NC-GANs have been used to augment wearable device datasets by simulating accelerometer and gyroscope noise. This augmentation improved the ability of models to distinguish subtle activities, such as walking and jogging, even in noisy environments [20].

NC-GANs extend the adversarial loss function to prioritize the fidelity of noise-induced features. Training on these datasets equips models to handle real-world sensor outputs, reducing errors caused by unpredictable noise.

2.5.3 Bridging Domain Discrepancies

Domain discrepancies, often caused by differences between training and deployment environments, pose significant challenges to model generalization. As mentioned, GANs address these discrepancies by aligning source and deployment domain distributions.

CycleGANs, a prominent GAN variant, achieve this alignment by learning mappings between unpaired source and target datasets. These networks employ two generators and two discriminators to iteratively refine mappings while preserving underlying data structures and adapting domain-specific characteristics. For instance, CycleGANs have been utilized to align laboratory-collected sensor data with outputs from real-world wearable devices. This alignment reduced error rates in activity recognition tasks by up to 15% [?].

2.6 Multimodal Data Integration in Activity Recognition

GANs have emerged as a powerful framework for addressing the complexities of multimodal data integration, particularly in activity recognition. By synthesizing and aligning heterogeneous data—such as video, audio, and various sensor signals—GANs create cohesive datasets that support accurate recognition tasks. They ensure *temporal coherence*, aligning events consistently across time, and *contextual coherence*, preserving meaningful relationships between data from different sources. For instance, temporal coherence ensures that video frames and sensor readings correspond to the same moment in time, while contextual coherence maintains logical connections between actions and environmental cues, enabling models to interpret diverse inputs effectively.

2.6.1 Addressing Temporal Misalignment and Data Heterogeneity

Integrating diverse data streams often involves challenges like temporal misalignment and discrepancies among data types. For example, accelerometer readings may update more frequently than gyroscope data or video frames, leading to inconsistencies that degrade system performance. Combining continuous time-series data with discrete image frames requires advanced methods to align features effectively.

GANs tackle this challenge by projecting multimodal data into a shared latent space, creating a unified representation where features from different sources are harmonized. What this means is that GANs transform data from various modalities into a common format, enabling models to integrate information without inconsistencies caused by differences in data structure. This approach captures both temporal dependencies and contextual relationships, forming a robust framework for accurate multimodal integration [25].

2.7 Handling Missing Modalities

Multimodal systems often face challenges related to missing modalities due to environmental factors or sensor failures. For example, video data may be unavailable in low-light conditions, or overlapping sounds may obscure audio inputs. GANs address this issue by generating synthetic data to fill in the gaps, ensuring continuity in multimodal systems.

Models like Noise-Conditional GANs (NC-GANs) and Multimodal GANs excel in handling incomplete datasets by conditioning the generation process on the available modalities. For instance, in the Multimodal Channel State Information-Based Activity Recognition (MCBAR) system, GANs generate spectral features from WiFi data, leading to a 12% improvement in recognition accuracy. This capability ensures that synthetic data complements existing data streams effectively, preserving overall system performance [5, 6].

2.7.1 Fusing Disparate Modalities

GANs excel at fusing disparate modalities by projecting their features into a shared latent space, minimizing inconsistencies between data types. The discriminator plays a crucial role in this process by ensuring that the generator's outputs align with the statistical properties of all input modalities. This fusion enhances interpretability and integration, allowing models to leverage complementary information from diverse data sources.

For example, in emotion recognition tasks, GANs unify text and facial expression data into a cohesive representation, reducing alignment errors and improving recognition accuracy. Similarly, in healthcare applications, GANs combine accelerometer data with audio signals, modeling complementary relationships between modalities to improve the detection of patient activities. These capabilities highlight GANs' effectiveness in synthesizing multimodal data for complex real-world scenarios [9, 10].

3 Challenges and Open Questions

Despite significant advancements, several open questions and challenges remain in utilizing GANs for activity recognition. First, training instability is a core challenge, as achieving a balance between the generator and discriminator during adversarial training is notoriously difficult. This instability often leads to mode collapse, where GANs produce repetitive outputs that fail to capture the diversity of real-world data distributions. For activity recognition, this lack of diversity undermines the quality and applicability of synthetic datasets. Additionally, GAN training is highly sensitive to hyperparameter settings, adding another layer of complexity to achieving stability [37, 38].

A critical question is whether GANs can effectively model rare or complex activities that involve intricate patterns or subtle gestures, particularly when training datasets are sparse. Current architectures often struggle to replicate such nuances, necessitating novel designs and training strategies to improve fidelity in these scenarios [39]. Furthermore, the quality and diversity of synthetic data generated by GANs remain significant limitations. While

GANs excel at generating general patterns, they frequently fail to represent subtle variations or complex environmental contexts. Reliable methods to evaluate and enhance synthetic data fidelity while maintaining variability to prevent overfitting are needed. Hybrid approaches, such as combining GANs with Variational Autoencoders or diffusion models, may offer promising solutions [40, 41].

The high computational cost associated with GAN training further exacerbates these challenges. Adversarial training demands substantial hardware resources and extended training times, particularly for large-scale or multimodal datasets. This creates a barrier to entry for smaller organizations or resource-constrained research teams. Optimizing GANs for real-time activity recognition systems, such as healthcare monitoring or surveillance, where low-latency processing is crucial, poses an additional challenge. Lightweight architectures and hardware acceleration may provide viable solutions for these applications [42, 43].

Another issue is the label scarcity often encountered in activity recognition datasets. GANs rely on labeled data for initial training, but for rare or complex activities, labeled data is inherently limited. This scarcity reduces GANs' ability to learn and replicate intricate activity patterns effectively. While data augmentation partially addresses this limitation, the effectiveness of GANs remains constrained by the diversity and availability of labeled examples [44].

Finally, all things ML, should lead to ethical and privacy concerns. GANs present significant challenges in the deployment of GANs. Although GANs is a tool to keep data private, it may inadvertently reproduce identifiable patterns from their training data, raising risks of data leakage and privacy violations. Additionally, the opaque nature of GAN-generated data raises ethical questions about its potential misuse, such as in creating deepfakes or deceptive synthetic datasets. These concerns necessitate the development of privacy-preserving GANs, incorporating techniques such as differential privacy and federated learning, to safeguard sensitive information while enabling effective training [45, 46].

4 Conclusion

GANs have effectively tackled challenges like data scarcity, domain adaptation, and multimodal integration, critical opportunities for growth remain.

Reducing high computational costs and addressing mode collapse are essential for broader adoption, particularly in resource-limited contexts. Advancing GAN architectures to capture rare or complex activities, such as nuanced gestures or dynamic environmental changes, is also crucial. Integrating lightweight GANs into real-time applications like healthcare monitoring and autonomous navigation could greatly enhance their practical value. Additionally, incorporating Explainable AI (XAI) frameworks into GANs would increase transparency and trust by elucidating how these models replicate data distributions, particularly in sensitive areas like medical and financial applications.

References

[1] J. K. Aggarwal and M. S. Ryoo, "Human Activity Analysis: A Review," *ACM Computing Surveys*, vol. 43, no. 3, pp. 1–43, 2011.

[2] A. Bulling, U. Blanke, and B. Schiele, "A Tutorial on Human Activity Recognition Using Body-Worn Inertial Sensors," *ACM Computing Surveys*, vol. 46, no. 3, pp. 1–33, 2014.

[3] L. Wang, Y. Xiong, Z. Wang, Y. Qiao, D. Lin, X. Tang, and L. Van Gool, "Temporal Segment Networks: Towards Good Practices for Deep Action Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 6, pp. 1530–1545, 2018.

[4] F. J. Ordóñez and D. Roggen, "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition," *Sensors*, vol. 16, no. 1, pp. 1–25, 2016.

[5] Z. Liu, J. Yin, and H. Wang, "Noise-Conditional Generative Adversarial Networks for Robust Multimodal Activity Recognition," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 10, pp. 10214–10223, 2021.

[6] R. Chen, Y. Zhang, and T. Huang, "Multimodal GANs for Activity Recognition with Missing Modalities," in *Proceedings of the IEEE International Conference on Multimodal Systems (ICMS)*, pp. 123–130, 2020.

[7] Zhao2020CycleGANWearable} H. Zhao, W. Li, Y. Chen, and T. Wang, "CycleGAN-Based Domain Adaptation for Activity Recognition Using Wearable Sensors," \emph{IEEE Sensors Journal}, vol. 20, no. 9, pp. 5015–5024, 2020.

[8] C. Little, M. J. Elliot, R. Allmendinger, and S. S. Samani, "GAN Architecture Diagram. From Generative Adversarial Networks for Synthetic Data Generation: A Comparative Study," Published under the Creative Commons Attribution 4.0 International License, 2021.

[9] J. Zhang, H. Li, and Y. Zhang, "Emotion Recognition Using GANs for Multimodal Data Fusion," *IEEE Transactions on Affective Computing*, vol. 12, no. 2, pp. 356–367, 2021.

[10] R. Chen, Y. Liu, and M. Wang, "Multimodal GANs for Healthcare Activity Recognition Using Audio and Accelerometer Data," in *Proceedings of the IEEE International Conference on Healthcare Informatics (ICHI)*, pp. 123–130, 2020.

[11] D. Ravi, C. Wong, B. Lo, and G. Z. Yang, "Deep Learning for Human Activity Recognition: A Resource Efficient Implementation on Low-Power Devices," in *Proceedings of the IEEE International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, 2016, pp. 71–76.

[12] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, 2010.

[13] Z. Zhang, Z. Zhang, C. Dong, and P. Xiao, "Multimodal Human Activity Recognition: A Comprehensive Survey," *IEEE Access*, vol. 9, pp. 81213–81238, 2021.

[14] R. Torkzadehmahani, P. Kairouz, and B. Paten, "DP-CGAN: Differentially Private Synthetic Data and Label Generation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2019.

[15] C. Shorten and T. M. Khoshgoftaar, "Data augmentation: A comprehensive survey," *Journal of Big Data*, vol. 6, no. 1, pp. 60, 2019.

[16] H. Fang, G. Wang, J. Dong, L. Yang, H. Liu, and S. Wang, "Generative Adversarial Networks for Medical Image Augmentation: A Path Towards Solving Imbalanced Data Issue," *IEEE Access*, vol. 8, pp. 59427–59440, 2020.

[17] G. Yin, J. Zhang, Z. Zhang, and Y. Wang, "Data augmentation for class imbalance in human activity recognition using GANs," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 11, pp. 4711–4721, 2020.

[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[19] T. Wang, Y. Li, X. Zhang, and H. Wang, "Generative Adversarial Networks for data augmentation in robotic hand gesture recognition," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2772–2780, 2021.

[20] W. Jiang, J. Yin, H. Wang, and Z. Zhang, "Noise-Conditioned Generative Adversarial Networks for Wearable Sensor Data Augmentation," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 9, pp. 3123–3134, 2019.

[21] C. Kumar, P. Gupta, A. Singh, and A. Sharma, "Automated Driving Scenario Generation Using GANs," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 12, pp. 5158–5167, 2020.

[22] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," in *Proceedings of the 34th International Conference on Machine Learning (ICML)*, vol. 70, pp. 214–223, 2017.

[23] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved Training of Wasserstein GANs," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, pp. 5767–5777, 2017.

[24] Y. Wang, Z. Lin, X. Chen, and L. Qi, "Improving Nighttime Activity Recognition Using GAN-Synthesized Data," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 3147–3156, 2020.

[25] Z. Liu, J. Li, M. Zhang, and Y. Chen, "Multimodal Data Integration Using GANs for Temporal Alignment in Human Activity Recognition," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 7, pp. 2983–2994, 2021.

[26] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive Growing of GANs for Improved Quality, Speed, and Stability," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2018.

[27] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "StyleGAN: A Style-Based Generator Architecture for Generative Adversarial Networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4401–4410, 2019.

[28] S. Sankaranarayanan, Y. Balaji, C. D. Castillo, and R. Chellappa, "Generate to Adapt: Aligning Domains Using Generative Adversarial Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8503–8512, 2018.

[29] M. Buades-Rubio, A. Lozano, and J. Mora, "Activity Recognition and Fall Detection Using Deep Learning: A Systematic Review," in *ArXiv Preprint arXiv:2401.11790*, 2024.

[30] W. Sultani, C. Chen, and M. Shah, "Real-World Anomaly Detection in Surveillance Videos," in *ArXiv Preprint arXiv:1801.04264*, 2018.

[31] Siemens, "Leveraging AI for Predictive Maintenance: The Future of Industrial Efficiency," in *Siemens Blog*, 2024. [Online]. Available: `https://blog.siemens.com/2024/08/leveraging-ai-for-predictive-maintenance-the-future-of-industrial-efficiency/`

[32] M. Hassan, A. Rahman, and T. Alam, "Adversarial AI for Cross-User and Inter-Domain Adaptation in Human Activity Recognition," in *PLOS ONE*, 2024.

[33] M. Hassan, A. Rahman, and T. Alam, "Human Activity Recognition Using Wireless Signals: Addressing Domain Adaptation Challenges," in *PLOS ONE*, 2024.

[34] J. Xaviar, K. Patel, and M. Singh, "Centaur: A Robust Multimodal Fusion Model for Human Activity Recognition," in *ArXiv Preprint arXiv:2303.04636*, 2023.

[35] Q. Zhang, W. Li, and X. Xu, "Deep Learning for Human Activity Recognition: A Comprehensive Review," in *Journal of Big Data*, vol. 8, no. 44, pp. 1–27, 2021.

[36] GeeksforGeeks, "Generative Adversarial Networks (GANs): An Overview," in *GeeksforGeeks Blog*, 2024. [Online]. Available: `https://www.geeksforgeeks.org/generative-adversarial-network-gan/`

[37] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," in *Proceedings of the 34th International Conference on Machine Learning (ICML)*, vol. 70, pp. 214–223, 2017.

[38] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Nets," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 27, pp. 2672–2680, 2014.

[39] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "StyleGAN: A Style-Based Generator Architecture for Generative Adversarial Networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 4401–4410, 2019.

[40] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2014.

[41] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 33, pp. 6840–6851, 2020.

[42] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *ArXiv Preprint arXiv:1704.04861*, 2017.

[43] H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, "EfficientNet: Scalable and Efficient Deep Neural Networks,"

in *Proceedings of the International Conference on Machine Learning (ICML)*, pp. 123–130, 2021.

[44] C. Shorten and T. M. Khoshgoftaar, "A Survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, pp. 60, 2019.

[45] R. Torkzadehmahani, P. Kairouz, and B. Paten, "DP-CGAN: Differentially Private Synthetic Data and Label Generation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2019.

[46] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, pp. 1273–1282, 2017.