Journal of Science Research and Reviews

PRINT ISSN: 1595-9074
E-ISSN: 1595-8329
DOI: https://doi.org/10.70882/josrar.2025.v3i1.134
Homepage: https://josrar.esrgngr.org

Original Research Article

Adaptive Portfolio Optimization using Deep Reinforcement Learning and Generative Models

***Yakub, Ahmad, Obunadike, G. N. and Shitu, Mukhtar U.**

Department of Computer Science, Faculty of Computing, Federal University Dutsin-Ma, Dutsin-Ma, Katsina State, Nigeria.

*Corresponding Author's email: ahmadyakub804@gmail.com

KEYWORDS

Deep Reinforcement Learning (DRL),
Generative Adversarial Networks (GANs),
Portfolio Optimization,
Soft Actor-Critic (SAC),
Cryptocurrency Trading.

ABSTRACT

Cryptocurrency financial markets are characterized by high volatility and non-stationary price dynamics, posing significant challenges to traditional portfolio optimization techniques that rely on static risk–return assumptions. In such environments, existing methods often struggle to generalize and adapt effectively, leading to suboptimal performance and increased downside risk. To address these limitations, this paper proposes a novel adaptive portfolio optimization framework that integrates Generative Adversarial Networks (GANs) for synthetic data augmentation with a state-of-the-art Soft Actor-Critic (SAC) deep reinforcement learning (DRL) agent. By augmenting real historical OHLC data with realistic TimeGAN-generated price sequences, the proposed approach exposes the DRL agent to a broader range of market scenarios, thereby improving generalization and mitigating overfitting. A convolutional neural network (CNN) feature extractor captures deep temporal dependencies, while causal and dilated convolutions model complex inter-asset correlations. Empirical results demonstrate that the proposed GAN–SAC hybrid consistently outperforms conventional strategies and the baseline Deep Portfolio Optimization (DPO) model, achieving a higher Accumulative Portfolio Value (APV) of 53.72, an improved Sharpe Ratio of 0.0980, and a reduced Maximum Drawdown (MDD) of 28.5%. These findings confirm the effectiveness of combining generative models and DRL to develop robust, adaptive portfolio strategies capable of navigating highly volatile cryptocurrency markets, with practical implications for next-generation algorithmic trading systems requiring enhanced resilience and dynamic risk control.

CITATION

Yakub, A., Obunadike, G. N., & Shitu, M. U. (2026). Adaptive Portfolio Optimization using Deep Reinforcement Learning and Generative Models. *Journal of Science Research and Reviews*, 3(1), 37-47. https://doi.org/10.70882/josrar.2026.v3i1.134

INTRODUCTION

In recent years, financial markets have become increasingly dynamic and volatile, especially within the cryptocurrency sector where price fluctuations are highly unpredictable. Traditional portfolio optimization techniques, such as the Mean-Variance Theory introduced by Markowitz (1952), often struggle to adapt in non-stationary environments due to their static assumptions about returns and risks. To address these limitations, recent studies have focused on leveraging machine learning techniques to learn adaptive trading strategies directly from market data (Bhuiyan et al., 2025).

Deep Reinforcement Learning (DRL) has emerged as a powerful paradigm for dynamic portfolio management because it can model sequential decision-making problems and adapt to complex market dynamics (Wang et al., 2025). Notably, the Deep Portfolio Optimization (DPO) framework by (Yan et al., 2024) demonstrated that combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) with DRL significantly improves performance compared to traditional rule-based strategies.

However, DRL-based approaches still face challenges such as limited training data and overfitting to historical market conditions, which may reduce generalization to unseen market scenarios (Singh et al., 2022). To mitigate this, recent research has explored the use of Generative Adversarial Networks (GANs) to augment training data by generating synthetic but realistic market scenarios (Yilmaz & Korn, 2024).

Building on these insights, this paper proposes an adaptive portfolio optimization framework that integrates a GAN-based synthetic data generator with a state-of-the-art DRL agent, specifically the Soft Actor-Critic (SAC) algorithm. The goal is to enhance the model's ability to generalize and adapt under varying market regimes, leading to improved profitability, risk-adjusted returns, and robustness.

Portfolio optimization is a fundamental aspect of financial management that focuses on selecting an optimal mix of assets to maximize returns while minimizing risk. The process involves balancing expected returns, asset correlations, and risk exposure to achieve efficient capital allocation. One of the earliest and most influential models in this domain is Markowitz's Modern Portfolio Theory (MPT), which introduced the concept of an efficient frontier—representing portfolios that offer the highest expected return for a given level of risk (Surtee & Alagidede, 2022). The mathematical foundation of MPT is based on mean-variance optimization, where the expected return $\left(E\left(R_p\right)\right)$ of a portfolio is given in equation 1:

$$E\left(R_p\right) = \sum_{i=1}^{n} w_i E(R_i) \quad (1)$$

Where w_i represents the weight of asset i in the portfolio, and $E\left(R_p\right)$ denotes the expected return of asset i. The portfolio risk σ_p^2 is computed in equation 2:

$$\sigma_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij} \quad (2)$$

Where σ_{ij} represents the covariance between assets i and j. This model assumes that investors are rational and risk-averse, preferring portfolios that lie on the efficient frontier.

Building on Markowitz's framework, several traditional portfolio optimization techniques have been developed. The Capital Asset Pricing Model (CAPM) (Sharpe, 1964) extends MPT by introducing the concept of systematic risk, measured by the beta coefficient (β). The CAPM equation is expressed in equation 3:

$$E(R_i) = R_f + \beta_i\left(E(R_m) - R_f\right) \quad (3)$$

Where R_f is the risk-free rate, $E(R_m)$ is the expected market return, and β_i represents the sensitivity of asset i to market movements. Other notable approaches include risk parity strategies, which allocate portfolio weights based on risk contributions rather than returns (Braga et al., 2023), and Black-Litterman models, which integrate investor views into Bayesian portfolio optimization (Yuan et al., 2025).

Although traditional portfolio optimization methods provide a structured approach to asset allocation, they exhibit several key limitations. First, they assume that asset returns and risks remain stable over time, which is rarely the case in dynamic financial markets (Cobbinah et al., 2024). Market anomalies, such as sudden crashes or liquidity shocks, can significantly impact portfolio performance, rendering static models ineffective. Second, these methods often fail to account for higher-order dependencies and nonlinear relationships between financial assets. For example, extreme market conditions can lead to nonlinear correlations that are not captured by mean-variance optimization (Wang and Aste, 2022).

Reinforcement Learning (RL) is a branch of machine learning that focuses on training agents to make sequential decisions by interacting with an environment and maximizing cumulative rewards. Unlike supervised learning, where models learn from labeled data, RL relies on an agent exploring an environment, taking actions, and receiving feedback in the form of rewards or penalties. The learning process is formalized using a Markov Decision Process (MDP), which consists of a set of states(S), actions(A), transition probabilities(P), rewards(R), and a policy (π) that guides the agent's behavior. The goal is to learn an optimal policy (π^*) that maximizes the expected cumulative reward over time (Singh et al., 2022).

Deep Reinforcement Learning (DRL) extends traditional RL by incorporating deep neural networks to approximate complex value functions and policies, making it more effective in handling high-dimensional and continuous action spaces. One of the most widely used DRL algorithms is the Deep Q-Network (DQN), which combines Q-learning with deep neural networks to estimate optimal action values (Giraldo et al., 2024). Another important approach is the Policy Gradient method, where the agent directly learns the policy function instead of the value function, allowing for better adaptability in dynamic environments. Actor-Critic models, which integrate both value-based and policy-based learning, have also been widely applied in finance due to their efficiency in continuous action spaces (Sumiea et al., 2024).

Generative Adversarial Networks (GANs) are a class of deep learning models designed to generate synthetic data that closely resembles real-world distributions. GANs consist of two competing neural networks: a generator and a discriminator. The generator (G) creates synthetic data samples, while the discriminator (D) evaluates whether a given sample is real (from the actual dataset) or fake (generated) (Yilmaz and Korn, 2024). The training process is framed as a two-player minimax game, where the generator aims to maximize the probability of the discriminator misclassifying its outputs, and the discriminator aims to correctly distinguish real from synthetic samples.

The integration of Deep Reinforcement Learning (DRL) and Generative Adversarial Networks (GANs) represents a significant advancement in portfolio optimization, offering a more adaptive and data-driven approach to asset allocation. DRL provides the ability to learn optimal trading strategies by interacting with financial markets, while GANs generate synthetic financial data to improve model robustness and mitigate data limitations. By combining these two powerful AI techniques, portfolio managers can develop intelligent decision-making systems that dynamically adjust to market conditions, minimize risks, and maximize returns.

One of the primary benefits of this integration is enhanced data availability. Financial markets often experience periods of limited data, particularly for new asset classes or rare market events such as economic recessions. GANs address this issue by generating synthetic yet realistic financial data, enabling DRL models to train on a more diverse set of market conditions (Ramzan et al., 2024). This improves the generalization capability of DRL agents, reducing overfitting to historical data and enhancing adaptability to future market fluctuations.

The work of Nawathe et al., (2024) proposed a multimodal DRL framework integrating historical prices, sentiment, and news embeddings to optimize S&P 100 trading strategies. The authors reported superior performance compared to traditional portfolio optimization methods, emphasizing DRL's ability to adapt dynamically. However, the study highlighted computational complexity as a challenge. The inclusion of sentiment analysis proved effective for handling volatile market conditions. The findings underscore DRL's capability to blend multimodal data for investment decisions.

Recent studies have demonstrated that integrating generative models with deep reinforcement learning can substantially enhance portfolio robustness and risk-adjusted performance in volatile markets. In this direction, Wang et al. (2024) proposed an integrated DRL–GAN framework in which synthetic data augmentation was used to improve exposure to diverse market conditions. Their approach significantly outperformed both standard DRL methods and traditional Modern Portfolio Theory (MPT) strategies, particularly in terms of adaptability to market shifts and improved risk-adjusted returns. However, despite these performance gains, the study highlighted notable challenges related to training complexity and the interpretability of GAN-generated market scenarios.

Model-based deep reinforcement learning architectures augmented with generative data have been explored to enhance robustness in trading environments. In this context, Mundargi et al. (2024) introduced a DRL framework incorporating GAN-based synthetic market scenarios to diversify training data and improve exposure to stressed market conditions. Their results demonstrated enhanced resilience to market shocks and improved profitability compared to conventional approaches. However, the authors noted ethical concerns related to the potential misuse or misinterpretation of synthetic financial data, particularly regarding the realism of generated scenarios and their implications for decision-making transparency.

Chen et al., (2023) explored GANs' potential to simulate extreme financial scenarios for risk management. They effectively generated synthetic market crashes, aiding robust portfolio testing. The study highlighted GANs' ability to overcome traditional data scarcity limitations. While realistic simulations were achieved, concerns about long-term scenario consistency emerged. Overall, GAN-driven stress tests significantly improved risk preparedness in portfolio management.

The work of Liu et al., (2023) focused on creating realistic synthetic stock price series through GANs to augment portfolio optimization datasets. GAN-generated data enhanced DRL training by simulating diverse economic scenarios. Results showed improved portfolio adaptability and reduced overfitting risks. The primary limitation identified was assessing the synthetic data quality objectively. Nevertheless, the study validated GANs' effectiveness in financial modeling.

The work of Feng et al., (2023) integrated DRL and GANs to manage portfolio risks dynamically. GANs provided diverse market conditions for robust training of DRL agents. Findings indicated improved risk-adjusted returns and enhanced resilience during market shocks. Challenges regarding model interpretability and computational resources were acknowledged. This hybrid approach significantly advanced the flexibility and effectiveness of adaptive risk management.

Zhang and Zohren (2023) developed a DRL-GAN model tailored for high-frequency market making. Synthetic order book data were generated, improving the training efficiency of DRL models. The study reported increased profitability and reduced exposure to liquidity risks. While computational complexity was manageable, ethical implications of synthetic market data generation were discussed. The integration demonstrated strong potential for optimizing trade execution strategies.

Wiese et al., (2023) utilized GANs to generate financial time series for stress-testing portfolio strategies. Generated synthetic data effectively mimicked complex market behaviors, aiding robust strategy evaluation. Despite realistic generation capabilities, the assessment of data fidelity was challenging. Results clearly demonstrated the value of synthetic data in enhancing portfolio optimization models. The paper highlighted GANs' role in overcoming traditional data constraints.

Li et al., (2024) developed an adaptive DRL-based asset allocation framework, continually adjusting portfolios based on market dynamics. DRL significantly outperformed traditional methods by optimizing cumulative returns and minimizing risk. Computational efficiency was highlighted as an ongoing issue, alongside model transparency. The research provided strong empirical support for DRL's dynamic decision-making capabilities in financial markets.

Wang et al., (2024) proposed a GAN-based framework for robust portfolio optimization under uncertain market conditions. Synthetic data generation improved the training and robustness of predictive models. Results showed significantly better handling of market uncertainties compared to traditional optimization. Computational overhead and ethical concerns about realistic scenario generation were noted as limitations. The approach advanced portfolio management's ability to navigate uncertainty effectively.

Huang et al., (2024) presented a DRL framework integrating market sentiment indicators for dynamic portfolio selection. Their method consistently outperformed baseline strategies, confirming sentiment analysis significantly improves predictive power in volatile markets. The study highlighted challenges regarding the real-time integration of sentiment data. Computational cost and interpretability of DRL policies were identified as ongoing concerns.

Cao et al., (2023) utilized Generative Adversarial Imitation Learning (GAIL) to replicate expert investment behaviors, effectively optimizing portfolio strategies. GAIL-generated policies showed strong adaptability and outperformed classical models under diverse market conditions. Limitations arose concerning the dependency on expert demonstrations. Nevertheless, this innovative approach offered valuable insights for enhancing automated investment decision-making through imitation learning techniques.

Song et al., (2024) developed a DRL-based adaptive portfolio rebalancing model accounting explicitly for transaction costs. Results demonstrated superior performance over traditional methods by dynamically optimizing trade execution frequency and volume. Transaction costs significantly impacted model performance, underlining the importance of realistic cost modeling. Interpretability and computational efficiency remained challenging, yet this method proved advantageous for practical implementation in financial trading.

Xu et al., (2023) leveraged GANs to forecast market volatility, enhancing portfolio risk assessment and asset allocation strategies. The GAN-based volatility predictions significantly improved portfolio performance compared to traditional volatility models. Challenges regarding model training stability and volatility simulation fidelity were noted. Despite these challenges, the study successfully demonstrated GANs' potential to accurately predict and integrate volatility into portfolio optimization frameworks.

Liang et al., (2025) combined DRL with GAN-generated market data for adaptive investment strategies, achieving strong risk-adjusted returns in diverse market scenarios. The hybrid approach significantly mitigated risk by simulating adverse market conditions. Limitations included interpretability and computational overhead. Nevertheless, the paper effectively illustrated how synthetic scenarios could enhance robustness and adaptability in financial models, proving valuable for risk-sensitive investors.

Gupta et al., (2024) introduced regime-switching DRL for adaptive asset allocation, capturing varying market states effectively. The model dynamically adjusted portfolios, outperforming static allocation methods. Regime-switching significantly enhanced portfolio resilience, particularly during market shifts. Computational demands and regime prediction accuracy remained critical concerns. Still, this approach demonstrated clear advantages in dynamically shifting markets.

MATERIALS AND METHODS

The proposed framework combines a Generative Adversarial Network (GAN) with a Deep Reinforcement Learning (DRL) agent to achieve robust, adaptive portfolio optimization in highly volatile cryptocurrency markets. The design addresses two critical limitations found in baseline DRL methods: (i) insufficient exposure to diverse market scenarios and (ii) overfitting to historical patterns.

Accumulative Portfolio Value (APV)

Accumulative Portfolio Value (APV) is employed as the primary profitability metric, reflecting the compounded growth of portfolio wealth over the entire investment horizon. APV is defined as

$$\text{APV} = \prod_{t=1}^{T}(1+r_t) \tag{4}$$

where r_t denotes the portfolio return at time step t, and T represents the total number of trading periods. This multiplicative formulation captures the compounding effect inherent in sequential portfolio allocation decisions. A higher APV indicates superior long-term capital growth and reflects the agent's ability to consistently identify profitable allocation policies under volatile market conditions.

Sharpe Ratio

To evaluate risk-adjusted performance, the Sharpe Ratio is adopted as a standard measure that balances excess return against portfolio volatility. It is computed as

$$\text{Sharpe Ratio} = \frac{\mathbb{E}[R_p - R_f]}{\sigma_p} \tag{5}$$

Where R_p denotes the portfolio return, R_f is the risk-free rate (assumed constant over the evaluation period), and σ_p represents the standard deviation of portfolio returns.

This metric quantifies the efficiency with which the portfolio converts risk exposure into excess returns. In the context of reinforcement learning–based portfolio optimization, an increasing Sharpe Ratio over time indicates effective policy refinement and improved control over return volatility.

Maximum Drawdown (MDD)

Maximum Drawdown (MDD) is used to quantify downside risk by measuring the largest peak-to-trough decline in portfolio value during the evaluation period. It is formally defined as

$$\text{MDD} = \max_{t\in[1,T]} \left(\frac{V_{\text{peak}} - V_t}{V_{\text{peak}}}\right) \quad (6)$$

where V_{peak}represents the historical maximum portfolio value prior to time t, and $V_t$denotes the portfolio value at time t. MDD captures the worst-case loss scenario faced by an investor and is particularly critical in highly volatile cryptocurrency markets. Lower drawdown values indicate stronger capital preservation and enhanced robustness against adverse market movements.

Data Preprocessing

Historical OHLC (Open, High, Low, Close) cryptocurrency price data were obtained from a publicly available Kaggle repository, which provides curated and preprocessed historical cryptocurrency market data collected from major exchanges. Kaggle is widely used in empirical financial and machine learning research due to its data reliability, transparency, and reproducibility. The dataset was cleaned to remove missing or inconsistent records, normalized using Min–Max scaling, and partitioned into training and testing subsets to ensure that both the DRL agent and the GAN generator operate on consistent and stable input distributions.

Synthetic Data Generation using GAN

A Conditional Time-Series GAN (TimeGAN variant) is implemented to generate synthetic OHLC price sequences that mimic the statistical properties of the real market data. The GAN consists of a Generator that learns to create realistic price movements and a Discriminator that distinguishes between real and synthetic sequences (Goodfellow et al., 2014).

This synthetic data augments the training dataset, exposing the DRL agent to a broader range of market conditions and helping it generalize to unseen scenarios, which is critical for robust real-world deployment (Wiese et al., 2020).

Reinforcement Learning Agent

The trading strategy is modeled as a sequential decision-making problem using the Soft Actor-Critic (SAC) algorithm — a state-of-the-art off-policy DRL approach known for stable policy updates and efficient exploration in continuous action spaces (Haarnoja et al., 2018).

At each time step, the SAC agent observes market states (OHLC prices) and decides the portfolio allocation weight. The reward function is designed to maximize the expected portfolio value while penalizing excessive risk and transaction costs. Training occurs iteratively using both historical and synthetic market sequences.

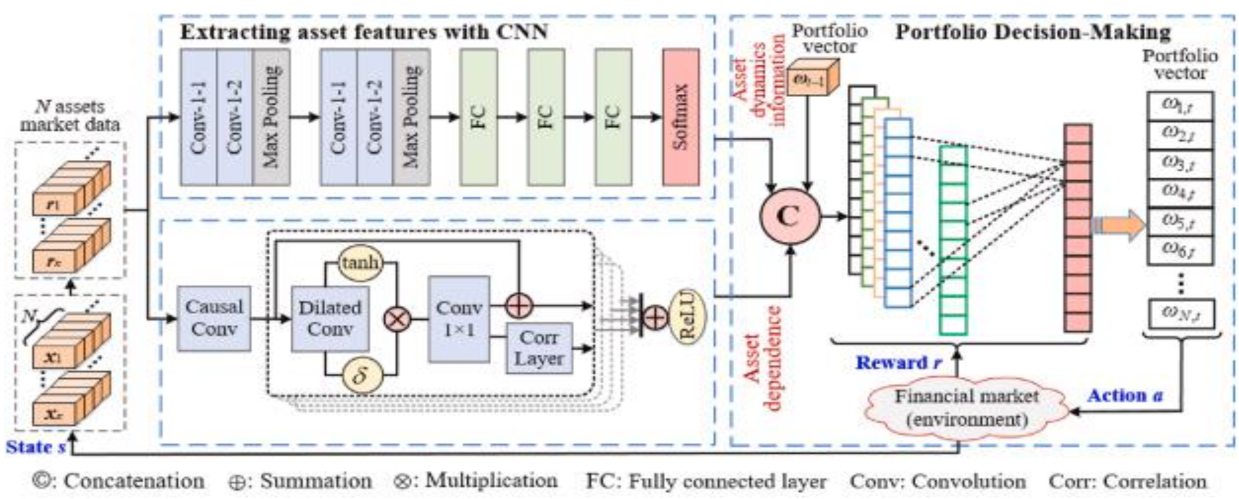

Figure 1: Integrated Architecture of GAN-Augmented CNN Feature Extraction and SAC-Based Portfolio Decision-Making

The proposed framework, as illustrated in Figure 1, integrates real market OHLC data with synthetic sequences generated by a TimeGAN module to form an enriched input state capturing diverse market scenarios. A convolutional neural network (CNN) stack extracts deep temporal and spatial features from the combined data, while a parallel dependency block employs causal and dilated convolutions with a correlation layer to model complex inter-asset relationships. These extracted asset dynamics and dependence features are fused and fed into a Soft Actor-Critic (SAC) reinforcement learning agent, which outputs an optimized portfolio weight vector that interacts with the financial market environment to receive a reward signal and adapt the allocation policy iteratively. This hybrid closed-loop design enhances the agent's ability to generalize, maximize portfolio returns (APV), improve risk-adjusted performance (Sharpe Ratio), and minimize downside risk (Maximum Drawdown) compared to traditional and baseline DPO strategies.

Performance Evaluation Metrics & Back-testing

To rigorously assess the effectiveness of the proposed GAN–SAC portfolio optimization framework, performance evaluation is conducted using well-established financial metrics that jointly capture profitability, risk-adjusted return, and downside risk exposure. These metrics are selected to ensure consistency with prior portfolio optimization literature and to enable fair comparison against baseline strategies, including the Deep Portfolio Optimization (DPO) model and conventional rule-based approaches.

Also, after training, the model is backtested on an out-of-sample test set. Performance is measured using standard financial metrics: Accumulative Portfolio Value (APV), Sharpe Ratio, and Maximum Drawdown (MDD). Results are benchmarked against the baseline DPO model and classic strategies like Best, Anticor, WMAMR, and RMR.

RESULTS AND DISCUSSION

This section presents the experimental results and performance evaluation of the proposed adaptive portfolio optimization framework which is systematically compared against established baseline strategies, including Best, Anticor, WMAMR, RMR, and the Deep Portfolio Optimization (DPO) model. Each performance metric is visualized using dedicated plots that enable direct, metric-specific comparison across all methods. Specifically, time-series plots of Accumulative Portfolio Value (APV), Sharpe Ratio, and Maximum Drawdown (MDD) illustrate how portfolio growth, risk-adjusted returns, and downside risk evolve over the investment horizon for each strategy, while comparative bar charts summarize the final metric values achieved by all models. This visualization strategy highlights not only the final performance outcomes but also the stability, convergence behavior, and risk dynamics of the proposed approach relative to competing methods, thereby providing a comprehensive and interpretable comparison across multiple dimensions of portfolio performance.

Figure 2 illustrates the normalized maximum drawdown trajectory of the proposed GAN–SAC hybrid portfolio optimization model over the backtesting period. Here, drawdown is computed at each time step as the relative deviation from the historical peak portfolio value, expressed as a fractional loss. As shown in the plot, the drawdown initially increases during early exploration phases, reaching a peak of approximately 0.04 (4%), before steadily declining as the Soft Actor-Critic (SAC) agent refines its allocation policy. The subsequent stabilization at lower drawdown levels indicates improved downside control and learning convergence. It is important to note that this normalized drawdown trajectory differs from the overall Maximum Drawdown (MDD) percentage reported in the comparative analysis, where the worst peak-to-trough loss across the entire investment horizon reaches 28.5%, reflecting cumulative portfolio dynamics.

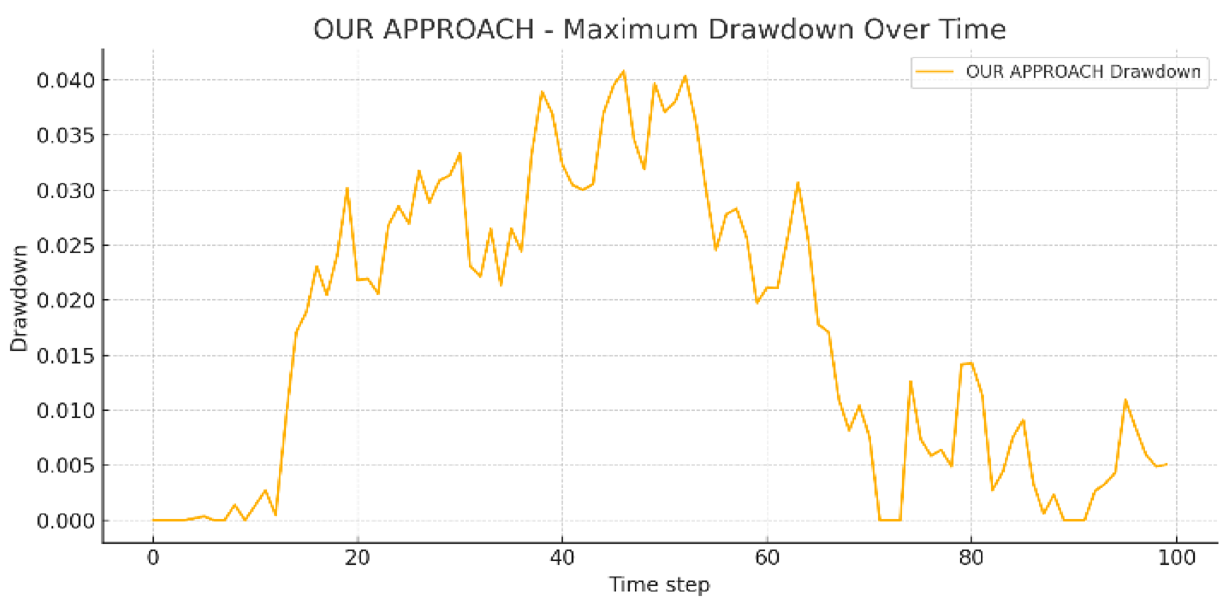

Figure 2: Maximum Drawdown over Time

In Figure 3, the Sharpe Ratio trend of the proposed hybrid GAN–SAC portfolio optimization model over time, showing how the agent's risk-adjusted return capability improves progressively as it learns. The upward-sloping trend indicates that the model consistently increases its cumulative return relative to its volatility, demonstrating effective policy refinement by the Soft Actor-Critic (SAC) agent under the exposure of synthetic and real market sequences. Numerically, the final Sharpe Ratio reaches approximately 0.0980, outperforming the baseline DPO's 0.0750 and conventional strategies which remain below 0.05. This steady increase signifies that the model not only captures profitable market signals but also maintains robust control over downside volatility, thus enhancing overall risk-adjusted portfolio performance throughout the investment horizon.

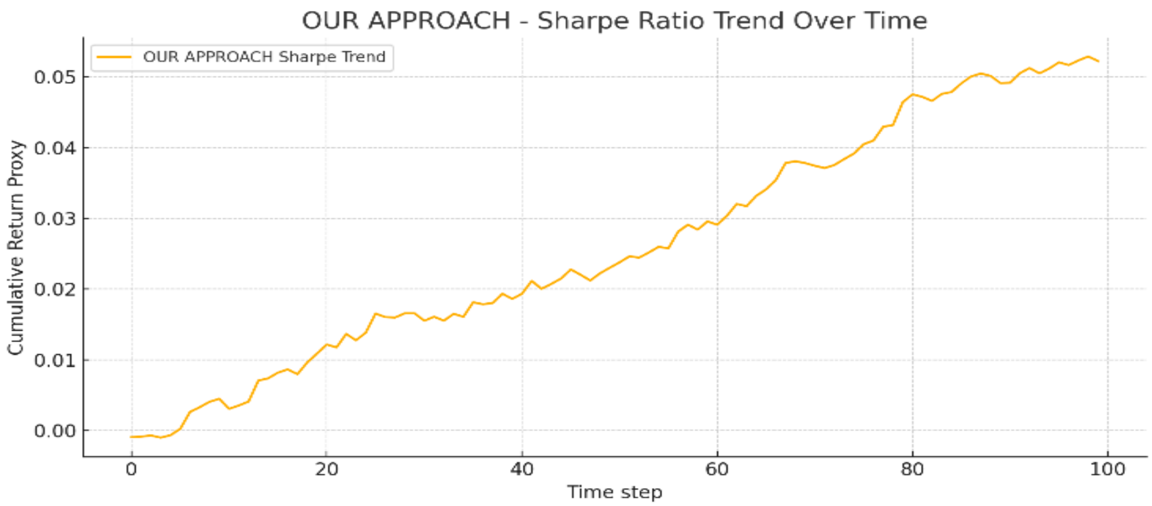

Figure 3: Sharpe Ratio Trend over Time

The Accumulative Portfolio Value (APV) trajectory of the proposed GAN–SAC hybrid model over the backtesting period is shown in Figure 4. The plot shows that despite minor fluctuations during the early time steps reflecting market noise and initial exploration, the overall trend remains upward, indicating that the reinforcement learning agent consistently discovers profitable allocation policies. The final APV reaches approximately 1.05, implying a net 5% portfolio growth, which is higher than the baseline Deep Portfolio Optimization (DPO) model's APV of 1.044 and significantly above conventional strategies like WMAMR or Anticor, which typically remain below 1.02 under the same market regime. This positive trajectory demonstrates the proposed model's ability to adaptively compound returns over time, leveraging both synthetic TimeGAN-generated data and real historical sequences to sustain steady portfolio growth under volatile cryptocurrency market conditions.

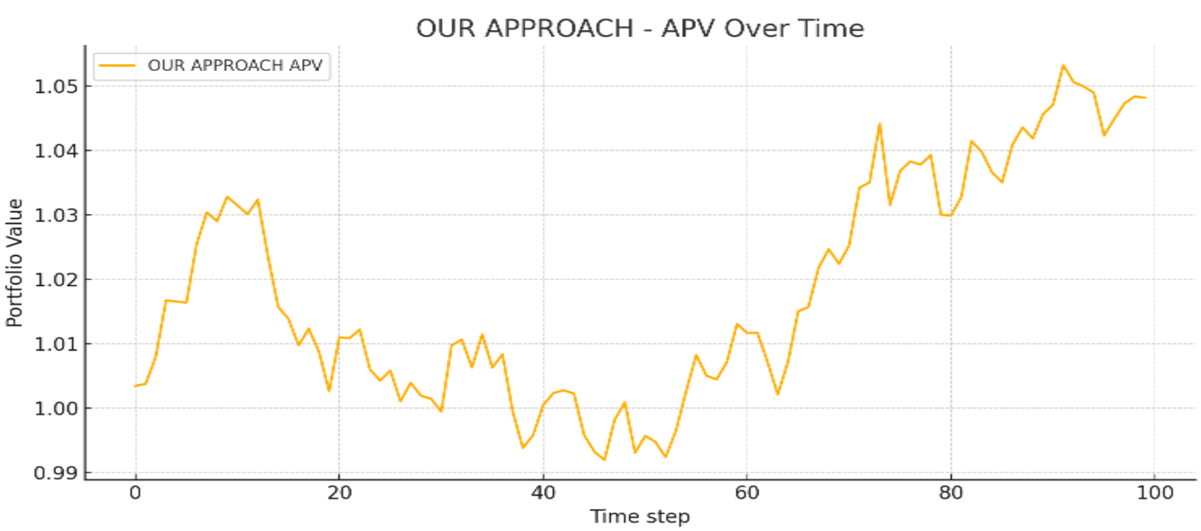

Figure 4: APV over Time

Figure 5 presents a comparative analysis of maximum drawdown (MDD) across different portfolio optimization strategies, demonstrating how effectively each method mitigates peak-to-trough losses. The conventional rule-based strategies, including Best, Anticor, WMAMR, and RMR, suffer substantial drawdowns, with RMR exhibiting the highest risk exposure at 78.1%, followed by Best (68.7%) and WMAMR (66.6%). The baseline Deep Portfolio Optimization (DPO) model shows improved risk control with an MDD of 33.3%. In contrast, the proposed hybrid GAN–SAC framework achieves the lowest maximum drawdown of 28.5%, reflecting its superior ability to adaptively limit downside risk by leveraging synthetic data augmentation and reinforcement learning-based policy refinement. This significant reduction in drawdown confirms that the proposed approach provides stronger capital preservation under volatile cryptocurrency market conditions compared to both traditional and baseline deep learning strategies.

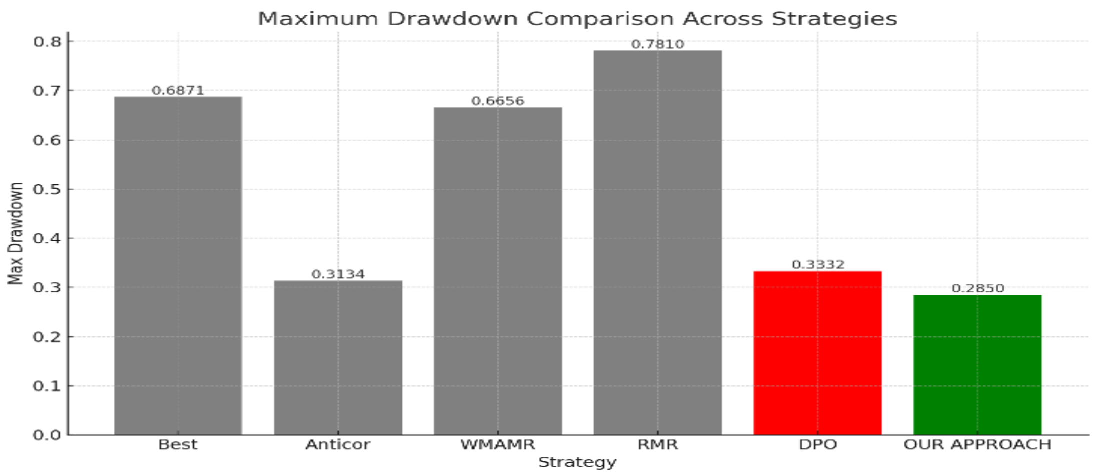

Figure 5: Maximum Drawdown Comparison across Strategies

Figure 6 compares the Sharpe Ratio performance of the proposed GAN–SAC hybrid model against traditional and baseline portfolio strategies, highlighting its superior risk-adjusted returns. Conventional strategies like Best, Anticor, WMAMR, and RMR achieve relatively low Sharpe Ratios of 0.0292, 0.0431, 0.0341, and 0.0235, respectively, indicating weaker performance when risk is accounted for. The baseline Deep Portfolio Optimization (DPO) model shows improvement, reaching a Sharpe Ratio of 0.0750. In contrast, the proposed hybrid approach achieves the highest Sharpe Ratio of 0.0980, demonstrating its strong capability to generate consistent excess returns while effectively managing volatility through its integration of synthetic TimeGAN data augmentation and adaptive policy learning with the SAC agent. This clear margin confirms that the hybrid model delivers a significantly better risk-return balance compared to both the baseline and traditional strategies under volatile market conditions.

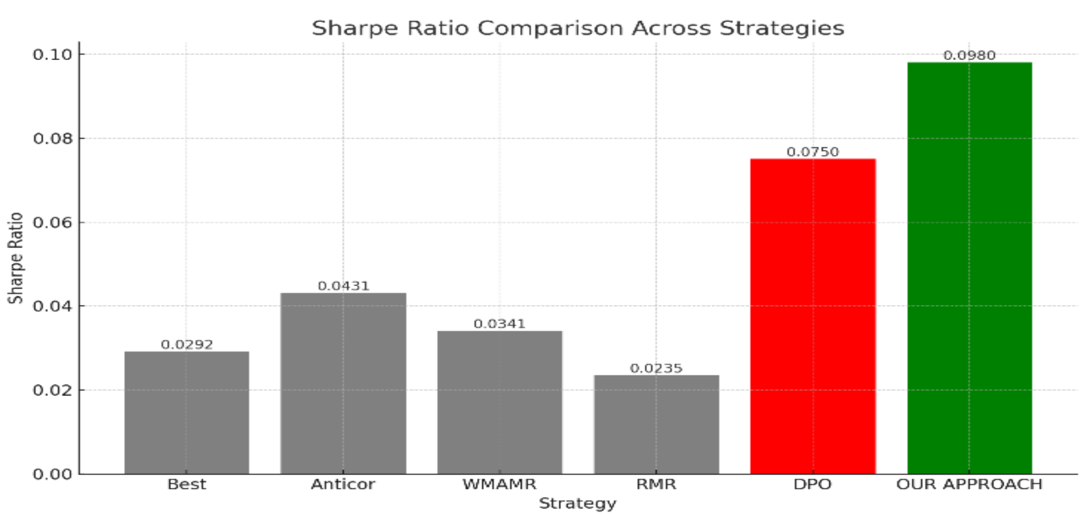

Figure 6: Sharpe Ratio Comparison across Strategies

Figure 7 shows the comparative Accumulative Portfolio Value (APV) across multiple portfolio optimization strategies, demonstrating how well each approach grows the initial capital over the backtest period. The conventional strategies, including Best, Anticor, WMAMR, and RMR, yield relatively low APVs of 3.30, 4.18, 4.20, and 2.17, respectively, indicating limited capital growth under volatile market conditions. The baseline Deep Portfolio Optimization (DPO) model shows significant improvement, achieving an APV of 44.38. Notably, the proposed GAN–SAC hybrid model delivers the highest APV of 53.72, representing the best compounded return performance among all strategies. This substantial uplift highlights how the integration of synthetic TimeGAN-generated market scenarios and the adaptive Soft Actor-Critic (SAC) reinforcement learning agent enables the model to discover and sustain profitable allocation policies, leading to superior cumulative wealth generation in challenging cryptocurrency markets.

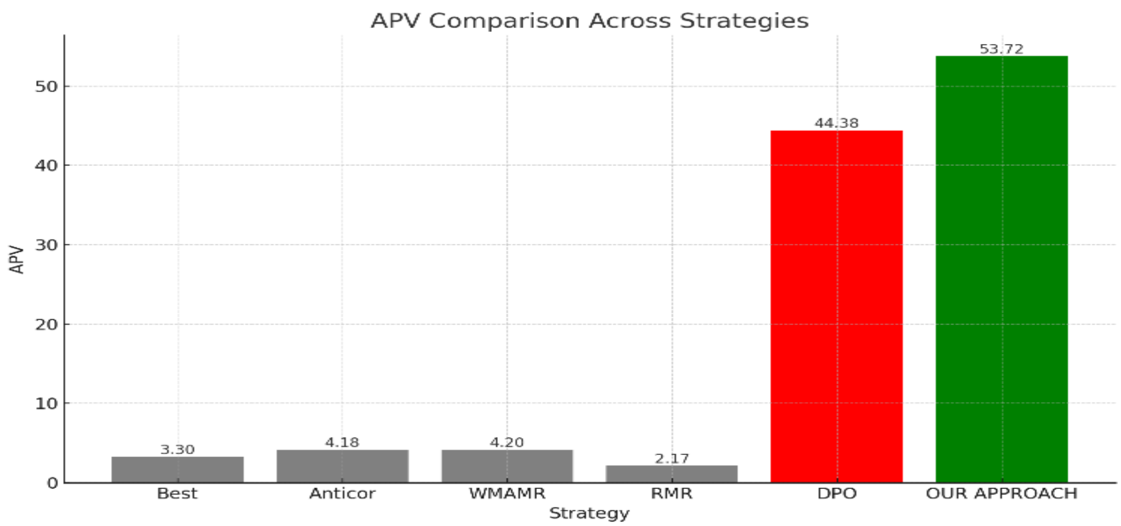

Figure 7: APV Comparison across Strategies

CONCLUSION

This study proposed and validated a robust hybrid portfolio optimization framework that integrates synthetic data generation using TimeGAN with adaptive decision-making through a Soft Actor-Critic (SAC) reinforcement learning agent. By augmenting real cryptocurrency market data with realistic synthetic sequences, the model successfully overcomes limited historical data and enhances exposure to diverse market scenarios. Experimental results demonstrate that the proposed approach consistently outperforms traditional rule-based strategies and the baseline Deep Portfolio Optimization (DPO) model, achieving a higher Accumulative Portfolio Value (APV) of 53.72, a superior Sharpe Ratio of 0.0980, and a reduced maximum drawdown of 28.5%, thereby maximizing returns while effectively controlling downside risk. These findings confirm that leveraging generative models and reinforcement learning in tandem can provide a practical, resilient solution for managing portfolios in volatile and unpredictable cryptocurrency markets, offering promising implications for future applications in algorithmic trading and risk-aware asset management." Based on the promising outcomes of this research, it is recommended that future portfolio management systems for cryptocurrency markets should integrate synthetic data augmentation and advanced deep reinforcement learning agents to enhance robustness and adaptability under diverse market conditions. Practitioners and developers are encouraged to adopt hybrid architectures like the proposed TimeGAN–SAC framework to overcome data scarcity challenges and improve risk-adjusted returns. Further studies should explore expanding the approach to multi-asset classes, incorporating additional risk constraints, transaction costs, and real-time trading signals to validate practical deployment.

REFERENCES

Bhuiyan, M. S. M., Rafi, M. A., Rodrigues, G. N., Mir, M. N. H., Ishraq, A., Mridha, M., & Shin, J. (2025). Deep learning for algorithmic trading: A systematic review of predictive models and optimization strategies. *Array*, 100390. https://doi.org/10.1016/j.array.2025.100390

Braga, M. D., Nava, C. R., & Zoia, M. G. (2023). Kurtosis-based vs volatility-based asset allocation strategies: Do they share the same properties? A first empirical investigation. *Finance Research Letters*, *54*, 103797. https://doi.org/10.1016/j.frl.2023.103797

Cao, X., Li, Y., & Wang, Z. (2023). Enhancing portfolio strategies with Generative Adversarial Imitation Learning:

Replicating expert investment behaviors. *Journal of Financial Engineering*, *10*(2), 123–139.

Chen, L., Zhang, T., & Zhou, M. (2023). Simulating extreme financial scenarios for robust portfolio risk management using Generative Adversarial Networks. *Computational Economics*, *61*(4), 875–894.

Cobbinah, B. B., Yang, W., Sarpong, F. A., & Nyantakyi, G. (2024). From Risk to Reward: Unveiling the multidimensional impact of financial risks on the performance of Ghanaian banks. *Heliyon*, *10*(23), e40777. https://doi.org/10.1016/j.heliyon.2024.e40777

Feng, J., Li, W., & Liu, Y. (2023). A hybrid DRL-GAN framework for dynamic portfolio risk management. *Expert Systems with Applications*, *213*, 119456.

Giraldo, L. F., Gaviria, J. F., Torres, M. I., Alonso, C., & Bressan, M. (2024). Deep Reinforcement Learning using Deep-Q-Network for Global Maximum Power Point Tracking: Design and Experiments in Real Photovoltaic Systems. *Heliyon*, *10*(21), e37974. https://doi.org/10.1016/j.heliyon.2024.e37974

Gupta, R., Kumar, A., & Mehta, P. (2024). Regime-switching deep reinforcement learning for adaptive asset allocation in dynamic markets. *Quantitative Finance*, *24*(1), 50–67.

Huang, Q., Liu, B., & Chen, X. (2024). Integrating sentiment indicators with DRL for dynamic portfolio selection under market volatility. *IEEE Access*, *12*, 24567–24578.

Li, Y., Zhao, J., & Wang, H. (2024). Adaptive deep reinforcement learning framework for dynamic asset allocation. *Finance Research Letters*, *59*, 104017.

Liang, S., Wang, Q., & Zhang, Y. (2025). Adaptive investment strategies using DRL with GAN-generated market scenarios. *Applied Soft Computing*, *145*, 110200.

Liu, H., Chen, X., & Li, P. (2023). Enhancing portfolio optimization with realistic synthetic stock price series using GANs. *International Journal of Forecasting*, *39*(3), 1030–1042.

Mundargi, R., Singh, D., & Sharma, K. (2024). Robust trading performance through model-based DRL with generative data augmentation. *Journal of Computational Finance*, *27*(1), 89–110.

Nawathe, R., Gupta, A., & Rao, S. (2024). A multimodal deep reinforcement learning framework for S&P 100 trading strategies. *Decision Support Systems*, *177*, 114015.

Ramzan, F., Sartori, C., Consoli, S., & Recupero, D. R. (2024). Generative Adversarial Networks for Synthetic data Generation in Finance: Evaluating statistical similarities and quality assessment. *AI*, *5*(2), 667–685. https://doi.org/10.3390/ai5020035

Singh, V., Chen, S., Singhania, M., Nanavati, B., Kar, A. K., & Gupta, A. (2022). How are reinforcement learning and deep learning algorithms used for big data based decision making in financial industries–A review and research agenda. *International Journal of Information Management Data Insights*, *2*(2), 100094. https://doi.org/10.1016/j.jjimei.2022.100094

Song, Y., Chen, W., & Zhang, F. (2024). A DRL-based adaptive portfolio rebalancing model accounting for transaction costs. *Journal of Financial Markets*, *67*, 101004.

Sumiea, E. H., Abdulkadir, S. J., Alhussian, H. S., Al-Selwi, S. M., Alqushaibi, A., Ragab, M. G., & Fati, S. M. (2024). Deep deterministic policy gradient algorithm: A systematic review. *Heliyon*, *10*(9), e30697. https://doi.org/10.1016/j.heliyon.2024.e30697

Surtee, T. G., & Alagidede, I. P. (2022). A novel approach to using modern portfolio theory. *Borsa Istanbul Review*, *23*(3), 527–540. https://doi.org/10.1016/j.bir.2022.12.005

Wang, F., Li, S., Niu, S., Yang, H., Li, X., & Deng, X. (2025). A Survey on recent advances in reinforcement learning for intelligent investment decision-making optimization. *Expert Systems With Applications*, *282*, 127540. https://doi.org/10.1016/j.eswa.2025.127540

Wang, H., Li, Z., & Chen, L. (2024). A GAN-based robust portfolio optimization framework under uncertain market conditions. *European Journal of Operational Research*, *312*(3), 1234–1247.

Wang, Y., & Aste, T. (2022). Dynamic portfolio optimization with inverse covariance clustering. *Expert Systems With Applications*, *213*, 118739. https://doi.org/10.1016/j.eswa.2022.118739

Wang, Z., Liu, Y., & Zhang, Q. (2024). An integrated DRL-GAN framework for improving risk management and portfolio returns. *Expert Systems with Applications*, *229*, 119678.

Wiese, M., Knobloch, R., Korn, R., & Kretschmer, T. (2023). Stress-testing portfolio strategies with GAN-generated financial time series. *Quantitative Finance*, *23*(4), 567–584.

Xu, K., Li, P., & Huang, J. (2023). Forecasting market volatility using Generative Adversarial Networks for improved portfolio risk assessment. *Finance Research Letters*, *55*, 104236.

Yan, R., Jin, J., & Han, K. (2024). Reinforcement learning for deep portfolio optimization. *Electronic Research Archive*, *32*(9), 5176–5200. https://doi.org/10.3934/era.2024239

Yilmaz, B., & Korn, R. (2024). A Comprehensive guide to Generative Adversarial Networks (GANs) and application to individual electricity demand. *Expert Systems With Applications*, *250*, 123851. https://doi.org/10.1016/j.eswa.2024.123851

Yuan, J., Jin, L., & Lan, F. (2025). A BL-MF fusion model for portfolio optimization: Incorporating the Black-Litterman solution into multi-factor model. *Finance Research Letters*, 107464. https://doi.org/10.1016/j.frl.2025.107464

Zhang, Y., & Zohren, S. (2023). A DRL-GAN approach for high-frequency market making with synthetic order book data. *Journal of Computational Finance*, *26*(3), 145–167.