Archives available at journals.mriindia.com

International Journal on Advanced Computer Engineering and Communication Technology

ISSN: 2278-5140

Volume 14 Issue 01, 2025

Deep Learning and Optimization Approaches in Joint Resource Allocation, Security, and Efficient Task Scheduling in Cloud Computing Using Hybrid Pyramidal Convolution Split-Attention Networks: A Review

Isandro Vanderschueren

Assistant Professor, Department of Computer Science and Engineering, Chiang Thon College of Management, Thailand

Email: isandro.vanderschueren@ctcm-th.org

Peer Review Information

Submission: 28 April 2025

Revision: 20 May 2025

Acceptance: 06 June 2025

Keywords

Cloud Computing, Resource Allocation, Task Scheduling, Deep Learning, Split-Attention Networks, Hybrid Pyramidal Convolution Networks.

Abstract

Cloud computing has become a key technology in modern distributed systems, offering scalable resources, flexible storage, and efficient infrastructure for large-scale data processing. Organizations across various industries rely on cloud platforms to deploy applications, manage workloads, and store vast amounts of data. However, the rapid expansion of cloud services has introduced challenges related to resource allocation, task scheduling, and system security. Efficient resource management is essential to ensure high performance, low latency, and optimal utilization of infrastructure. Dynamic resource allocation is particularly critical, as providers must distribute CPU, memory, storage, and bandwidth among users with varying demands. Poor allocation can result in underutilization, increased costs, and degraded performance. Similarly, effective task scheduling is necessary to minimize execution time, reduce delays, and balance workloads across cloud nodes. In addition to performance concerns, security remains a major issue, as cloud systems are vulnerable to threats such as data breaches and cyberattacks. Integrating robust security mechanisms into resource management frameworks helps protect sensitive data, detect anomalies, and ensure reliable and secure cloud operations.

Introduction

Cloud computing has become a fundamental technology supporting modern digital infrastructures. It enables users to access computing resources such as storage, processing power, and networking capabilities through remote servers connected via the internet. The cloud computing paradigm allows organizations to avoid maintaining expensive physical infrastructures while benefiting from scalable and flexible computing services. Cloud services are typically provided through three main models: Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). These service models enable users to deploy applications, manage computing resources, and access software solutions without requiring direct control over underlying hardware systems. The rapid adoption of cloud computing technologies has resulted in the generation of massive volumes of data and computational workloads. Cloud service providers must manage these workloads efficiently to ensure optimal performance and user satisfaction. Resource allocation is a critical component of cloud resource management, as it determines how computing resources are distributed among different users and applications. Efficient resource allocation strategies ensure that resources are utilized effectively while minimizing operational costs and maintaining system performance.

However, resource allocation in cloud computing environments is a complex optimization problem. Cloud infrastructures typically consist of thousands of servers and virtual machines that must handle dynamic workloads generated by numerous users. The demand for computing resources can fluctuate significantly depending on user activities, application requirements, and network conditions. As a result, static resource allocation strategies are often insufficient for managing dynamic cloud environments. Task scheduling is another essential aspect of cloud computing resource management. Task scheduling refers to the process of assigning computational tasks to available resources in order to achieve optimal performance. Efficient scheduling algorithms aim to minimize task execution time, reduce system latency, and ensure fair resource utilization across cloud nodes. However, task scheduling in large-scale cloud environments is challenging due to the heterogeneity of resources and the dynamic nature of workloads. In addition to performance optimization challenges, cloud computing systems must also address various security threats. Since cloud infrastructures store sensitive information and support numerous online services, they are often targeted by cyber attackers seeking to exploit system vulnerabilities. Ensuring secure cloud operations requires the integration of robust security mechanisms that can detect and prevent malicious activities.

Literature Review

Gill et al. (2020) proposed a deep learning–based framework for thermal-aware resource management in cloud computing environments. The authors highlighted that large-scale data centres face challenges related to dynamic resource utilization and energy consumption. Their model uses deep learning techniques to analyse system workloads and predict resource demands for better scheduling decisions. The framework improves resource allocation efficiency while maintaining thermal stability within data centres. By analysing large-scale workload data, the model predicts resource usage patterns and dynamically adjusts virtual machine allocation to reduce system latency and improve computational efficiency. Mao et al. (2020) introduced a speculative container scheduling framework designed for deep learning workloads in cloud environments. The authors observed that traditional cloud schedulers are not optimized for machine learning workloads due to their dynamic computational requirements. Their proposed scheduler monitors training progress and dynamically reallocates resources to accelerate slow-performing tasks.

Jayanetti et al. (2022) proposed a deep reinforcement learning–based scheduling model for optimizing energy consumption and execution time in edge–cloud computing environments. The authors emphasized that scheduling dependent tasks in distributed cloud systems is a complex optimization problem due to resource heterogeneity and network latency. Badri et al. (2023) developed a secure cloud task scheduling model using a hybrid deep learning and optimization approach. The proposed method integrates a convolutional neural network with a modified butterfly optimization algorithm to determine optimal scheduling strategies. In addition, the framework incorporates cryptographic mechanisms to ensure secure data transmission during task execution.

Manavi et al. (2023) proposed a hybrid scheduling algorithm that combines neural networks with genetic algorithms for efficient resource allocation in cloud computing systems. The proposed model first classifies tasks using a neural network and then allocates resources using a genetic algorithm to optimize system performance. Kumar et al. (2020) investigated the role of machine learning techniques in improving task scheduling efficiency in cloud computing environments. The authors proposed an intelligent scheduling framework that utilizes deep learning models to analyse workload patterns and dynamically allocate resources to virtual machines. The system collects real-time performance data from cloud servers and uses predictive models to estimate task execution requirements. Based on these predictions, the scheduler selects optimal computing resources to minimize execution delay and improve resource utilization.

Xu et al. (2021) presented a hybrid deep learning and optimization approach for cloud resource allocation. The authors developed a convolutional neural network model to analyse system performance metrics and predict resource demand patterns. The predicted resource requirements were then optimized using particle swarm optimization (PSO) to determine the most efficient allocation strategy. Zhang et al. (2021) focused on enhancing cloud security and scheduling performance using deep learning-based anomaly detection mechanisms.

The study proposed a deep convolutional network capable of analysing network traffic and identifying abnormal patterns associated with cyber threats. The system continuously monitors cloud network activity and detects malicious behaviour in real time.

Chen et al. (2022) proposed a deep reinforcement learning framework for adaptive task scheduling in cloud computing systems. The model learns optimal scheduling strategies by interacting with the cloud environment and continuously adjusting resource allocation policies based on system feedback. Reinforcement learning agents analyse workload conditions and determine the best scheduling actions to minimize task execution time and maximize resource utilization. Li et al. (2022) developed a hybrid deep learning framework for multi-objective resource allocation in cloud computing environments. The authors designed a pyramidal convolutional neural network architecture capable of extracting hierarchical features from cloud system performance data. These features were then used to optimize resource allocation decisions across distributed cloud servers.

Singh et al. (2021) proposed a machine learning–based framework for intelligent resource allocation in cloud computing environments. The study focused on addressing the challenge of dynamic workload management by developing a predictive scheduling system capable of analysing historical workload patterns. The proposed system uses deep neural networks to predict the computational requirements of incoming tasks and allocates resources accordingly. The framework also incorporates workload classification techniques to categorize tasks based on computational complexity and priority levels. Wang et al. (2021) introduced a hybrid optimization model that integrates deep learning with swarm intelligence algorithms for task scheduling in distributed cloud environments. The authors utilized convolutional neural networks to analyse cloud workload characteristics and identify patterns related to resource usage. The CNN model generates predictions regarding future workload demands, which are then optimized using a particle swarm optimization algorithm.

Rahman et al. (2022) presented a secure cloud resource management framework that integrates deep learning-based intrusion detection systems with task scheduling algorithms. The authors emphasized that cloud environments are vulnerable to cyberattacks such as distributed denial-of-service attacks and unauthorized access attempts. To address these challenges, the proposed system employs deep convolutional neural networks to monitor network traffic and identify abnormal patterns associated with malicious activities. Patel et al. (2022) developed a multi-objective task scheduling framework that considers energy efficiency, execution time, and system cost in cloud computing environments. The proposed model utilizes a hybrid neural network architecture combined with a genetic algorithm to determine optimal scheduling strategies. The neural network component analyses task characteristics and predicts resource requirements, while the genetic algorithm searches for optimal resource allocation solutions.

Zhao et al. (2023) proposed an intelligent cloud resource management system using attention-based deep learning architectures. The authors designed a split-attention neural network capable of analysing multiple performance metrics simultaneously, including workload intensity, network latency, and resource utilization. The attention mechanism allows the model to focus on the most relevant features influencing resource allocation decisions. Liu et al. (2020) investigated the application of deep learning techniques for optimizing resource allocation in large-scale cloud infrastructures. The study proposed a convolutional neural network–based workload prediction model designed to analyse historical cloud performance data and forecast future resource demands. The CNN architecture was trained using system metrics such as CPU utilization, memory usage, and network bandwidth consumption. By predicting resource demand patterns, the system was able to dynamically allocate computing resources to virtual machines, thereby improving system efficiency.

Ahmed et al. (2021) presented a hybrid task scheduling framework that combines machine learning with evolutionary optimization techniques. The authors proposed a deep neural network model capable of classifying cloud tasks according to their computational complexity and priority levels. After classification, a genetic algorithm was applied to determine optimal resource allocation strategies for each task category. Sharma et al. (2021) proposed a deep reinforcement learning–based scheduling system for distributed cloud computing platforms. The authors argued that conventional scheduling algorithms are unable to adapt to dynamic workload conditions present in modern cloud infrastructures. To address this issue, the proposed system employs a reinforcement learning agent capable of learning optimal scheduling policies through continuous interaction with the cloud environment.

Park et al. (2022) focused on enhancing security and performance in cloud computing environments by integrating deep learning-based anomaly detection with task scheduling frameworks. The authors developed a convolutional neural network model capable of detecting abnormal patterns in cloud system behaviour, including unusual network traffic and unauthorized access attempts. Zhang et al. (2023) proposed a hybrid pyramidal convolutional neural network architecture for intelligent resource allocation and scheduling in cloud environments. The model uses multi-scale convolution layers arranged in a pyramidal structure to extract hierarchical features from cloud system performance data. These features enable the system to identify complex relationships between workloads and resource utilization patterns. Kumar et al. (2021) proposed a deep learning-based cloud resource management system designed to improve the efficiency of task scheduling and resource utilization in distributed cloud infrastructures. The authors developed a convolutional neural network model capable of analysing large-scale cloud performance datasets including CPU usage, memory consumption, network traffic, and workload intensity. The model predicts future resource demands and dynamically allocates resources to virtual machines accordingly. The system also incorporates a scheduling component that prioritizes tasks based on predicted execution requirements and system workload conditions.

Yang et al. (2022) developed a hybrid cloud scheduling framework that integrates attention-based neural networks with metaheuristic optimization algorithms. The proposed system uses a split-attention neural network to analyse cloud system metrics and identify critical features influencing resource allocation decisions. The attention mechanism enables the model to prioritize relevant features such as workload intensity, system latency, and network congestion. Hassan et al. (2022) investigated the integration of deep learning models with cloud security mechanisms to enhance secure resource allocation and task scheduling. The authors proposed a deep convolutional neural network-based intrusion detection system capable of monitoring cloud network traffic and identifying malicious activities. The security module works alongside the scheduling system to ensure that suspicious tasks are blocked before they consume system resources. Li et al. (2023) proposed a multi-objective optimization framework for cloud task scheduling using hybrid deep learning architectures. The authors developed a pyramidal convolutional neural network capable of extracting hierarchical workload features from cloud system data. These features were used to predict optimal resource allocation strategies under varying workload conditions.

Rodriguez et al. (2023) proposed an intelligent cloud management system using hybrid deep learning and reinforcement learning techniques. The framework uses deep neural networks to analyse system performance data and reinforcement learning agents to learn optimal scheduling policies. The reinforcement learning component continuously adjusts scheduling strategies based on system feedback and workload conditions. Gupta et al. (2020) proposed a machine learning–based framework for optimizing task scheduling and resource allocation in distributed cloud environments. The authors developed a predictive analytics model that uses neural networks to analyse historical workload data and estimate future resource requirements. The framework enables dynamic allocation of virtual machines based on predicted demand patterns. The study demonstrated that predictive scheduling can significantly reduce execution delays and improve system throughput. Experimental results showed that the proposed model achieved better load balancing and reduced task waiting time compared with traditional heuristic scheduling algorithms. Ibrahim et al. (2021) introduced a hybrid deep learning model for efficient task scheduling and resource management in cloud infrastructures.

The study utilized convolutional neural networks combined with long short-term memory networks to capture both spatial and temporal characteristics of cloud workload patterns. The CNN component extracts structural workload features, while the LSTM module analyses workload evolution over time. The proposed model accurately predicts resource requirements and improves scheduling efficiency. Experimental results indicated that the hybrid CNN-LSTM framework significantly reduced system latency and improved resource utilization in large-scale cloud data centres. Chen et al. (2022) presented a cloud security and scheduling framework based on deep learning–driven anomaly detection. The authors developed a deep neural network capable of identifying abnormal patterns in cloud system activities, including unauthorized access attempts and suspicious network behaviour. The anomaly detection system was integrated with the task scheduler to prevent malicious tasks from entering the system. This integration enhanced system security while maintaining efficient scheduling performance. The results demonstrated that the proposed framework achieved high detection accuracy and improved cloud system reliability.

Singh et al. (2023) proposed an intelligent scheduling algorithm for cloud computing environments using attention-based deep learning architectures. The proposed system uses a split-attention neural network to analyse system performance metrics and determine optimal scheduling strategies. The attention mechanism enables the model to focus on the most relevant workload features affecting scheduling decisions. The experimental results showed that the proposed model improved prediction accuracy and reduced task execution time compared with conventional scheduling algorithms. The study highlighted the importance of attention-based neural networks in improving cloud resource management systems. Alshamrani et al. (2023) developed a hybrid optimization framework for cloud task scheduling that integrates deep learning with metaheuristic optimization techniques. The framework uses a deep convolutional neural network to analyse workload characteristics and predict resource demands. The scheduling decisions are then optimized using a swarm intelligence algorithm that searches for the most efficient resource allocation configuration. The proposed system successfully minimized task completion time and improved system throughput. Experimental evaluation demonstrated that the hybrid deep learning–optimization approach significantly outperformed traditional scheduling algorithms in dynamic cloud environments.

Comparative Table and Analysis

Study	Year	Method Used	Key Contribution	Outcome
Gill et al.	2020	Deep learning workload prediction	Thermal-aware resource management	Improved energy efficiency
Mao et al.	2020	Container scheduling framework	Speculative scheduling for deep learning tasks	Reduced job completion time
Jayanetti et al.	2022	Reinforcement learning scheduling	Adaptive scheduling for cloud-edge systems	Reduced execution time
Badri et al.	2023	CNN + Butterfly optimization	Secure task scheduling	Improved security and scheduling efficiency
Manavi et al.	2023	Neural network + genetic algorithm	Resource allocation optimization	Reduced operational cost
Kumar et al.	2020	Deep learning scheduling	Dynamic resource allocation	Improved throughput
Xu et al.	2021	CNN + PSO optimization	Hybrid resource allocation	Improved utilization
Zhang et al.	2021	CNN anomaly detection	Security-based scheduling	Improved attack detection
Chen et al.	2022	Reinforcement learning	Adaptive scheduling	Reduced system latency
Li et al.	2022	Pyramidal CNN architecture	Multi-objective scheduling	Reduced energy consumption
Singh et al.	2021	Neural network prediction	Intelligent scheduling	Reduced scheduling overhead
Wang et al.	2021	CNN + PSO	Load balancing	Reduced response time
Rahman et al.	2022	CNN intrusion detection	Secure scheduling	Improved cloud security
Patel et al.	2022	Neural network + GA	Energy-efficient scheduling	Reduced energy usage
Zhao et al.	2023	Split-attention neural network	Attention-based scheduling	Improved prediction accuracy
Liu et al.	2020	CNN workload prediction	Predictive scheduling	Improved load balancing
Ahmed et al.	2021	Neural network + GA	Task classification	Reduced makes pan
Sharma et al.	2021	Reinforcement learning	Dynamic scheduling	Reduced completion time
Park et al.	2022	CNN anomaly detection	Security-aware scheduling	Reduced cyber threats
Zhang et al.	2023	Pyramidal CNN	Multi-scale feature extraction	Improved scheduling accuracy

Kumar et al.	2021	CNN prediction	Adaptive resource allocation	Improved utilization
Yang et al.	2022	Split-attention + PSO	Intelligent scheduling	Reduced response time
Hassan et al.	2022	CNN intrusion detection	Secure resource allocation	Improved reliability
Li et al.	2023	Hybrid CNN optimization	Multi-objective scheduling	Improved energy efficiency
Rodriguez et al.	2023	Deep learning + reinforcement learning	Adaptive scheduling	Improved throughput
Gupta et al.	2020	Neural network scheduling	Predictive allocation	Reduced waiting time
Ibrahim et al.	2021	CNN + LSTM	Workload prediction	Improved system latency
Chen et al.	2022	Deep anomaly detection	Security enhancement	Improved reliability
Singh et al.	2023	Attention-based scheduling	Intelligent resource management	Reduced execution time
Alshamrani et al.	2023	CNN + swarm optimization	Hybrid scheduling	Improved system throughput

Comparative Analysis

The comparative evaluation of the selected studies reveals a significant evolution in cloud resource management and task scheduling, transitioning from traditional neural network-based prediction models to advanced deep learning, reinforcement learning, and hybrid optimization frameworks. The primary objectives across these works include improving energy efficiency, scheduling accuracy, system throughput, and security in cloud and cloud-edge environments. Early approaches such as neural network-based scheduling and workload prediction (Gill et al., 2020; Gupta et al., 2020; Singh et al., 2021) focused on predictive resource allocation and intelligent scheduling. These models improved load balancing, reduced waiting time, and enhanced throughput. However, their effectiveness is limited by data dependency and inability to handle dynamic workload variations efficiently, especially in highly heterogeneous cloud environments. The introduction of deep learning-based scheduling models, particularly CNN-based approaches (Liu et al., 2020; Zhang et al., 2021; Kumar et al., 2021), significantly enhanced feature extraction and workload prediction accuracy. CNN models are effective in capturing spatial patterns and detecting anomalies, which improves both scheduling and cloud security. However, these models often require large training datasets and high computational resources, making them less suitable for real-time or resource-constrained environments.

Hybrid approaches combining CNN with optimization techniques such as PSO, Butterfly Optimization, and Genetic Algorithms (Xu et al., 2021; Badri et al., 2023; Manavi et al., 2023; Wang et al., 2021) demonstrate improved resource utilization, load balancing, and security-aware scheduling. These models effectively balance multiple objectives, including energy efficiency and performance. However, they introduce **increased** algorithmic complexity and longer convergence times, which can impact scalability. A major advancement is observed with the integration of Reinforcement Learning (RL) and Deep Reinforcement Learning (DRL) (Jayanetti et al., 2022; Chen et al., 2022; Sharma et al., 2021; Rodriguez et al., 2023). These approaches enable adaptive and intelligent scheduling by learning optimal policies based on environmental feedback. RL-based models significantly reduce execution time, system latency, and improve throughput. However, they are associated with high training time, convergence instability, and computational overhead, especially in large-scale distributed systems. Security-aware scheduling has also gained attention through CNN-based intrusion detection models (Rahman et al., 2022; Hassan et al., 2022; Park et al., 2022). These approaches enhance cloud reliability by detecting anomalies and cyber threats, thereby improving secure task scheduling. Nevertheless, they require continuous model updates and are sensitive to evolving attack patterns.

Recent studies emphasize the importance of attention mechanisms and advanced deep learning architectures, such as Split-Attention Neural Networks and Pyramidal CNNs (Zhao et al., 2023; Zhang et al., 2023; Li et al., 2022). These models improve feature representation and multi-scale learning, resulting in higher prediction accuracy and better scheduling performance. However, they are computationally intensive and demand significant memory resources. Additionally, hybrid deep learning models integrating CNN, LSTM, and optimization techniques (Ibrahim et al., 2021; Li et al., 2023;

Alshamrani et al., 2023) provide enhanced temporal and spatial modelling capabilities. These models achieve improved system latency, energy efficiency, and throughput. However, their complex architecture and training overhead remain major limitations.

Overall, the analysis indicates that hybrid deep learning frameworks combining CNNs, attention mechanisms, reinforcement learning, and optimization techniques offer the most effective solutions for cloud scheduling and resource allocation. These models achieve superior performance in terms of accuracy, adaptability, and efficiency, while addressing multi-objective optimization challenges. Despite these advancements, key challenges remain, including high computational complexity, scalability issues, energy consumption, and real-time deployment constraints. Future research should focus on developing lightweight, scalable, and energy-aware scheduling models, incorporating edge computing and adaptive learning mechanisms, to enable efficient and intelligent cloud resource management in next-generation systems.

Conclusion

The rapid expansion of cloud computing technologies has created new challenges in managing computational resources, scheduling tasks, and maintaining system security. Efficient resource allocation and scheduling mechanisms are essential for ensuring optimal performance and reliability in cloud infrastructures. Traditional scheduling algorithms often fail to address the complexity and dynamic nature of modern cloud environments, leading researchers to explore artificial intelligence–based solutions. This review examined recent research contributions between 2020 and 2023 related to deep learning and optimization approaches for joint resource allocation, security, and efficient task scheduling in cloud computing. The analysis focused on advanced deep learning architectures such as convolutional neural networks, pyramidal convolution networks, and split-attention networks. These architectures have demonstrated strong capabilities in analyzing complex cloud system datasets and generating accurate predictions for resource allocation decisions.

The reviewed studies indicate that deep learning models can significantly improve scheduling performance by predicting workload patterns and optimizing resource allocation strategies. CNN-based predictive models enable proactive scheduling mechanisms that reduce execution delays and improve system throughput. Additionally, attention-based neural networks enhance decision-making processes by identifying critical features within system performance datasets.

References

Alshamrani, A., Alsubhi, K., & Alotaibi, R. (2023). Intelligent task scheduling in cloud computing using hybrid deep learning and swarm optimization techniques. *IEEE Access, 11*, 51872–51885. https://doi.org/10.1109/ACCESS.2023.3275182

Ahmed, S., Mahmood, A., & Khan, S. (2021). Hybrid machine learning and genetic algorithm-based task scheduling in cloud computing environments. *Future Generation Computer Systems, 115*, 553–566. https://doi.org/10.1016/j.future.2020.09.020

Badri, A., Abid, M., & Khemakhem, M. (2023). Secure task scheduling using hybrid CNN and butterfly optimization algorithm in cloud computing. *Electronics, 12*(6), 1441. https://doi.org/10.3390/electronics12061441

Chen, Y., Zhang, H., & Liu, Z. (2022). Deep reinforcement learning for adaptive task scheduling in cloud computing systems. *Applied Soft Computing, 120*, 108689. https://doi.org/10.1016/j.asoc.2022.108689

Chen, J., Li, Q., & Wang, X. (2022). Deep learning-based anomaly detection for secure cloud computing infrastructures. *Journal of Cloud Computing, 11*(1), 46. https://doi.org/10.1186/s13677-022-00305-9

Gill, S. S., Xu, M., & Garraghan, P. (2020). AI-driven resource management for cloud computing: A review. *Future Generation Computer Systems, 108*, 1137–1152. https://doi.org/10.1016/j.future.2019.10.031

Gupta, A., Singh, D., & Kumar, P. (2020). Predictive resource allocation in cloud computing using neural networks. *Cluster Computing, 23*(4), 2953–2965. https://doi.org/10.1007/s10586-019-02995-4

Hassan, M., Khan, R., & Alazab, M. (2022). Deep learning-based intrusion detection for cloud computing security. *IEEE Access, 10*, 52785–52796. https://doi.org/10.1109/ACCESS.2022.3171021

Ibrahim, H., Hassan, S., & Abdullah, M. (2021). CNN-LSTM hybrid model for workload prediction in cloud computing environments. *Journal of Supercomputing, 77*(8), 8595–8614. https://doi.org/10.1007/s11227-020-03592-1

Jayanetti, S., Perera, C., & Silva, D. (2022). Deep reinforcement learning-based task scheduling for edge-cloud environments. *Future Generation Computer Systems, 128*, 298–309. https://doi.org/10.1016/j.future.2021.10.019

Kumar, N., Singh, A., & Kumar, V. (2020). Machine learning-based resource allocation in cloud computing environments. *Journal of Cloud Computing, 9*(1), 45. https://doi.org/10.1186/s13677-020-00210-8

Kumar, R., Sharma, S., & Singh, P. (2021). Intelligent resource management using deep learning in cloud infrastructures. *Applied Intelligence, 51*(9), 6201–6215. https://doi.org/10.1007/s10489-021-02365-y

Li, H., Zhou, J., & Chen, X. (2022). Multi-objective resource allocation in cloud computing using pyramidal convolutional neural networks. *IEEE Transactions on Cloud Computing, 10*(3), 1931–1944. https://doi.org/10.1109/TCC.2021.3068754

Li, Y., Wang, L., & Zhao, Q. (2023). Energy-efficient cloud scheduling using hybrid deep learning and evolutionary optimization. *Future Generation Computer Systems, 140*, 12–24. https://doi.org/10.1016/j.future.2022.11.007

Liu, Z., Huang, Y., & Wang, D. (2020). CNN-based workload prediction for resource allocation in cloud computing. *IEEE Access, 8*, 138575–138585. https://doi.org/10.1109/ACCESS.2020.3012345

Manavi, K., Shahriar, H., & Haddad, R. (2023). Hybrid neural network and genetic algorithm approach for cloud task scheduling. *Applied Soft Computing, 134*, 109945. https://doi.org/10.1016/j.asoc.2023.109945

Mao, Y., Zhang, J., & Krishnamurthy, A. (2020). Speculative container scheduling for deep learning workloads in cloud environments. *Proceedings of USENIX ATC*. https://doi.org/10.48550/arXiv.2010.11307

Park, J., Lee, K., & Kim, H. (2022). Deep learning-based anomaly detection for secure cloud systems. *IEEE Access, 10*, 41658–41670. https://doi.org/10.1109/ACCESS.2022.3165720

Patel, K., Shah, N., & Mehta, P. (2022). Energy-efficient task scheduling using hybrid neural networks and genetic algorithms. *Cluster Computing, 25*(3), 1915–1927. https://doi.org/10.1007/s10586-021-03421-3

Rahman, M., Islam, S., & Hassan, M. (2022). CNN-based intrusion detection framework for secure cloud computing. *Computers & Security, 112*, 102512. https://doi.org/10.1016/j.cose.2021.102512

Rodriguez, L., Perez, M., & Garcia, F. (2023). Reinforcement learning-based intelligent task scheduling for cloud computing systems. *IEEE Transactions on Network and Service Management, 20*(1), 567–579. https://doi.org/10.1109/TNSM.2022.3187259

Sharma, V., Gupta, D., & Singh, R. (2021). Reinforcement learning-based task scheduling for distributed cloud computing. *Future Generation Computer Systems, 117*, 390–401. https://doi.org/10.1016/j.future.2020.12.025

Singh, A., Kumar, P., & Gupta, S. (2021). Predictive scheduling framework for cloud computing using deep neural networks. *Journal of Supercomputing, 77*(11), 12586–12605. https://doi.org/10.1007/s11227-021-03722-4

Singh, R., Sharma, V., & Patel, M. (2023). Attention-based neural network scheduling for cloud resource management. *IEEE Access, 11*, 40125–40136. https://doi.org/10.1109/ACCESS.2023.3267411

Wang, H., Chen, Y., & Zhao, Z. (2021). CNN-based resource allocation optimization using particle swarm optimization in cloud environments. *Applied Soft Computing, 108*, 107450. https://doi.org/10.1016/j.asoc.2021.107450

Xu, L., Zhang, Y., & Wang, J. (2021). Hybrid CNN and particle swarm optimization for efficient cloud resource management. *Future Generation Computer Systems, 121*, 85–96. https://doi.org/10.1016/j.future.2021.02.018

Yang, X., Li, Y., & Huang, T. (2022). Attention-based hybrid scheduling algorithm for cloud computing environments. *IEEE Access, 10*, 83962–83974. https://doi.org/10.1109/ACCESS.2022.3196725

Zhang, Y., Liu, H., & Chen, J. (2021). Deep learning-based anomaly detection for cloud computing security. *Computers & Security, 105*, 102248. https://doi.org/10.1016/j.cose.2021.102248