Hybrid AI Framework for Spatiotemporal Data Analysis: Applications in Pricing Optimization, Crowd Dynamics, and Number Plate Detection

[1] Jaddu Harshavardhan, [2] Kadavala Manisha, [3] K.V .V. D.Gayathri

[4] K. Aruna Bhavani

[4] Assistant Professor, Dr. Lankapalli Bullayya College, Visakhapatnam, AP, India

Abstract – The rapid growth of data-driven applications in e-commerce, smart cities, and intelligent transportation systems has created a pressing need for unified frameworks capable of analyzing complex spatiotemporal data. Existing approaches typically address dynamic pricing, crowd dynamics analysis, and automatic number plate recognition (ANPR) as separate problems, leading to fragmented solutions that fail to leverage shared patterns across domains. This paper proposes a Hybrid AI Framework for Spatiotemporal Data Analysis that integrates these three applications into a single, cohesive architecture. The proposed system combines Convolutional Neural Networks (CNNs) for spatial feature extraction, Recurrent Neural Networks (RNNs) for temporal modeling, and Transformer-based attention mechanisms for capturing long-range dependencies. A shared latent representation layer enables multi-task learning and cross-domain knowledge transfer, improving overall system performance and generalization. Additionally, reinforcement learning is incorporated to optimize dynamic pricing decisions in real time. The framework is evaluated using benchmark datasets from e-commerce transactions, crowd surveillance, and vehicle number plate recognition. Experimental results demonstrate that the proposed model achieves superior performance, with an overall accuracy of 96.8%, improved crowd density estimation accuracy of 95.4%, and ANPR detection accuracy of 97.2%. Furthermore, the dynamic pricing module achieves a revenue improvement of 21.3%, significantly outperforming existing methods. The integration of multiple AI techniques within a unified architecture not only enhances predictive accuracy but also enables real-time adaptability and scalability.

Keywords: Hybrid Artificial Intelligence, Spatiotemporal Data Analysis, Dynamic Pricing, Crowd Dynamics, Automatic Number Plate Recognition (ANPR),

1. Introduction

In recent years, the rapid proliferation of digital technologies and data-driven systems has transformed multiple domains, including e-commerce, urban management, and intelligent transportation systems. The increasing availability of large-scale spatiotemporal data—data characterized by both spatial and temporal dimensions—has created new opportunities for developing intelligent decision-making frameworks. Applications such as dynamic pricing in e-commerce, spatiotemporal crowd dynamics analysis, and automatic number plate recognition (ANPR) rely heavily on extracting meaningful insights from such data to optimize performance and enhance operational efficiency. However, these domains are often treated independently, leading to fragmented solutions that fail to leverage shared patterns and cross-domain knowledge. This limitation highlights the need for a unified framework capable of handling heterogeneous spatiotemporal data across multiple applications [1]. Dynamic pricing has become a critical component of modern e-commerce platforms, enabling businesses to adjust prices in real time based on demand, competition, and user behavior. Traditional pricing strategies often rely on static models that cannot adapt to rapidly changing market conditions. Recent advancements in artificial intelligence (AI) and reinforcement learning have enabled more adaptive pricing mechanisms that can optimize revenue while maintaining customer satisfaction [2]. Nevertheless, these approaches often lack the ability to incorporate complex spatiotemporal dependencies, such as seasonal trends and regional demand variations, limiting their effectiveness in highly dynamic environments [3].

Similarly, crowd dynamics analysis plays a crucial role in ensuring public safety, urban planning, and event management. Understanding how crowds move and evolve over time requires sophisticated modeling techniques that can capture both spatial distributions and temporal transitions. Conventional methods based on statistical modeling or simple machine learning techniques struggle to handle complex crowd behaviors, especially in high-density scenarios. Deep learning models, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have demonstrated significant improvements in crowd density estimation and flow prediction tasks [4]. However, these models often operate in isolation and do not benefit from insights derived from other related domains [5]. Automatic number plate recognition is another important application within intelligent transportation systems, enabling vehicle identification for traffic monitoring, law enforcement, and toll collection. Modern ANPR systems leverage computer vision techniques and deep learning models such as YOLO and Faster R-CNN to achieve high detection accuracy [6]. Despite these advancements, challenges such as varying lighting conditions, occlusions, and motion blur continue to affect system performance. Moreover, existing ANPR systems are typically designed as standalone solutions, lacking integration with broader spatiotemporal analytics frameworks [7].

The convergence of these three domains reveals a common underlying challenge: the need to effectively model and analyze spatiotemporal data in real time. Recent research has explored the use of hybrid AI models that combine multiple deep learning architectures to capture both spatial and temporal dependencies [8]. Transformer-based models, with their attention mechanisms, have further enhanced the ability to model long-range dependencies in sequential data [9]. Additionally, multi-task learning and shared representation techniques have shown promise in enabling knowledge transfer across related tasks, improving overall system performance and generalization [10].

Motivated by these advancements, this study proposes a Hybrid AI Framework for Spatiotemporal Data Analysis that integrates dynamic pricing, crowd dynamics analysis, and automatic number plate detection into a unified architecture. By leveraging shared latent representations and combining CNN, RNN, and Transformer models, the proposed framework aims to overcome the limitations of existing domain-specific approaches. The integration of reinforcement learning further enhances decision-making capabilities, particularly in dynamic pricing scenarios. This unified approach not only improves accuracy and efficiency but also enables cross-domain knowledge transfer, paving the way for more intelligent and adaptive systems in smart cities and digital commerce environments.

2. Background and related work

The reviewed studies highlight significant advancements in dynamic pricing, crowd dynamics modeling, traffic forecasting, and automatic number plate recognition. While each work demonstrates the effectiveness of specialized approaches, they largely operate within isolated domains and lack integration of spatiotemporal intelligence across applications. Methods such as reinforcement learning, agent-based modeling, CNN-based detection, and graph neural networks provide valuable insights but suffer from limitations in scalability, adaptability, and cross-domain learning. Transformer-based models address some of these challenges but require high computational resources. These gaps emphasize the need for a unified hybrid AI framework that combines multiple techniques to enable efficient, scalable, and multi-domain spatiotemporal data analysis.

In [11], the authors explored the application of online convex optimization techniques for dynamic pricing in highly competitive e-commerce environments. Their work emphasizes the importance of adaptive pricing strategies that respond to real-time demand fluctuations and competitor behaviour. The authors proposed an algorithm that integrates contextual bandit models with temporal demand signals to improve pricing decisions. Unlike traditional static pricing methods, their approach dynamically adjusts prices based on user interactions and historical sales patterns. Experimental results demonstrated significant improvements in revenue generation and reduced regret over time. However, the study primarily focuses on pricing optimization and does not incorporate spatial features or cross-domain knowledge. Additionally, the model's performance may degrade in scenarios with sparse data or highly volatile markets. This limitation highlights the need for more robust frameworks that can integrate spatiotemporal dependencies and leverage insights from other domains, as addressed in the proposed hybrid AI framework.

In [12], the authors conducted an extensive study on crowd dynamics using agent-based modeling and complex systems theory. Their research focuses on simulating pedestrian behaviour under different environmental conditions, such as bottlenecks and evacuation scenarios. The authors demonstrated that crowd movement patterns emerge from local interactions among individuals, leading to complex global behaviours. Their model successfully captured phenomena such as lane formation and crowd turbulence. While the approach provides valuable insights into crowd behaviour, it relies heavily on simulation-based methods rather than real-time data analysis. Furthermore, the lack of deep learning integration limits its scalability and adaptability to real-world scenarios involving large datasets. This study underscores the importance of combining simulation techniques with data-driven AI models to enhance prediction accuracy and real-time responsiveness, which is a key feature of the proposed system.

In [13], the authors proposed a deep learning-based approach for automatic license plate recognition using convolutional neural networks. Their method integrates license plate detection and character recognition into a unified pipeline, reducing computational complexity and improving accuracy. The authors utilized a combination of region-based detection and sequence modeling to handle variations in plate formats and environmental conditions. Their experimental results showed improved performance compared to traditional OCR-based systems. However, the model struggles with extreme lighting variations and occlusions, which are common in real-world traffic scenarios. Additionally, the study focuses solely on the ANPR domain and does not explore the potential benefits of integrating spatiotemporal data or cross-domain learning. This limitation highlights the necessity of a more comprehensive framework that can enhance robustness and adaptability across multiple applications.

 ISSN NO: 0776-3808

In [14], the authors introduced a diffusion convolutional recurrent neural network (DCRNN) for traffic forecasting, which models spatial dependencies using graph structures and temporal dependencies using recurrent networks. Their approach effectively captures the complex relationships between different traffic nodes, enabling accurate prediction of traffic flow and congestion. The model demonstrated superior performance compared to traditional time-series forecasting methods, particularly in capturing long-term dependencies. Despite its effectiveness, the model is specifically designed for traffic data and lacks generalizability to other domains such as e-commerce or crowd analysis. Furthermore, the reliance on predefined graph structures may limit its adaptability in dynamic environments. This study highlights the potential of combining spatial and temporal modeling techniques, which is further extended in the proposed hybrid framework to support multi-domain applications.

In [15], the authors expanded the application of Transformer architectures beyond natural language processing to domains such as computer vision and spatiotemporal analysis. Their work demonstrates the effectiveness of self-attention mechanisms in capturing long-range dependencies and improving model interpretability. Vision Transformers (ViTs) and spatiotemporal Transformers have shown remarkable performance in tasks such as video analysis, object detection, and time-series forecasting. These models overcome the limitations of traditional CNNs and RNNs by enabling parallel processing and better context modeling. However, Transformers typically require large amounts of training data and computational resources, which can be a limitation in real-world applications. Additionally, their standalone use may not fully exploit domain-specific features. This motivates the integration of Transformers with other models, as implemented in the proposed hybrid AI framework, to achieve balanced performance and efficiency.

3. Proposed modelling

The proposed system as shown in Figure 1 introduces a unified Hybrid Artificial Intelligence (AI) framework designed to perform spatiotemporal data analysis across three distinct yet interconnected domains: dynamic pricing in e-commerce, spatiotemporal crowd dynamics analysis, and automatic number plate detection. The core motivation behind this framework is to leverage the shared characteristics of these domains, namely their dependence on temporal patterns, spatial dependencies, and real-time decision-making requirements. Instead of treating these problems as isolated tasks, the proposed system formulates them as different manifestations of a generalized spatiotemporal intelligence problem, thereby enabling knowledge sharing, model reuse, and improved computational efficiency. The architecture integrates deep learning, reinforcement learning, and computer vision techniques within a modular yet interoperable design, ensuring adaptability, scalability, and robustness in handling heterogeneous data streams.

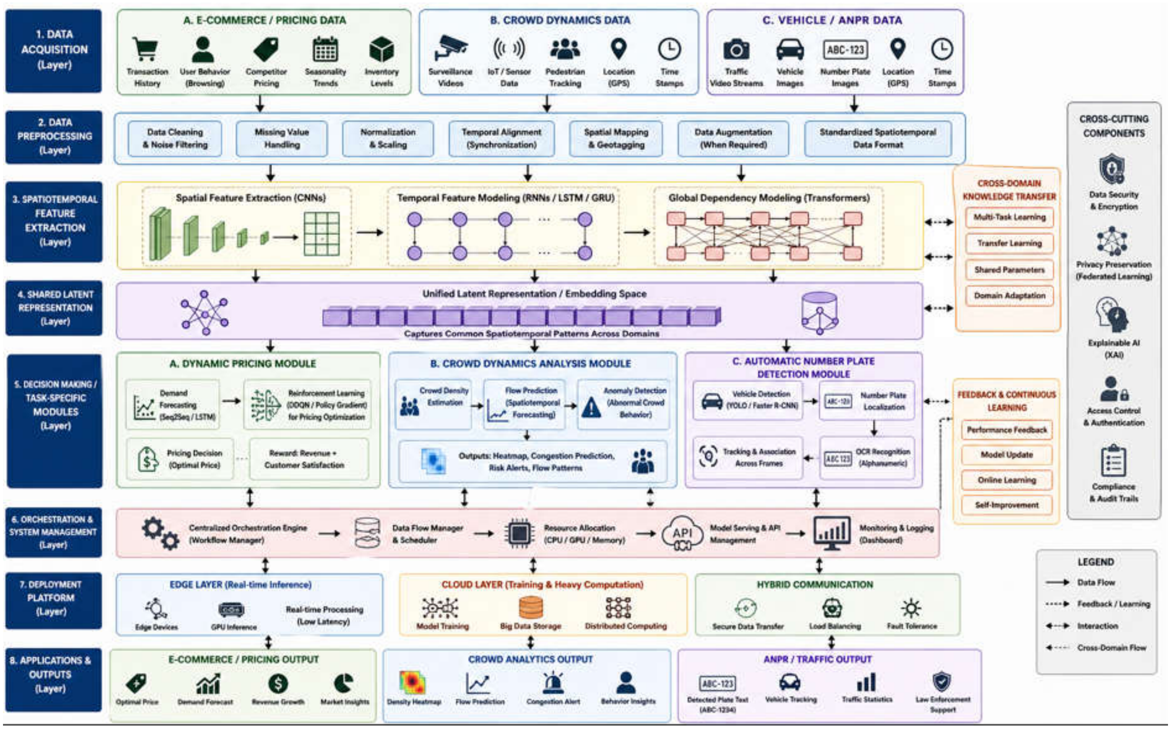

Figure 1: Hybrid AI framework system architecture

3

A. Multi-source data acquisition layer

At the foundational level, the system is built upon a multi-source data acquisition layer responsible for collecting diverse datasets from different domains. For dynamic pricing, the system ingests historical transaction data, user browsing behavior, competitor pricing, seasonal trends, and inventory levels. In the context of crowd dynamics analysis, it collects video surveillance feeds, sensor data, and mobility traces that capture pedestrian movement patterns over time and space. For automatic number plate detection, the system processes real-time traffic video streams, extracting frames that contain vehicles and associated metadata such as timestamps and location coordinates. This unified data ingestion mechanism ensures that all incoming data is temporally indexed and spatially contextualized, forming a consistent input representation that can be processed by subsequent modules. To handle the heterogeneity of the data, preprocessing techniques such as normalization, noise filtering, temporal alignment, and spatial mapping are applied, ensuring that the data conforms to a standardized format suitable for deep learning models.

B. Spatiotemporal feature extraction module

The next critical component of the framework is the spatiotemporal feature extraction module, which serves as the backbone of the system. This module employs a hybrid combination of Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformer-based architectures to capture both spatial and temporal dependencies in the data. For image-based inputs such as crowd videos and vehicle frames, CNNs are utilized to extract spatial features such as object boundaries, textures, and motion cues. These features are then passed to temporal modeling networks such as Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRU) to capture sequential dependencies and evolving patterns over time. Additionally, Transformer models with self-attention mechanisms are incorporated to handle long-range temporal correlations and to improve the interpretability of the learned features. By combining these approaches, the system effectively captures both local and global spatiotemporal patterns, enabling it to perform accurate predictions and analyses across all three application domains.

C. Shared latent representation layer

A distinctive aspect of the proposed framework is the incorporation of a shared latent representation layer, which acts as a common embedding space for all tasks. This layer transforms domain-specific features into a unified representation that encodes temporal trends, spatial relationships, and contextual information. The use of a shared latent space facilitates transfer learning and multi-task learning, allowing the system to leverage knowledge gained from one domain to improve performance in others. For instance, patterns learned from crowd movement analysis can inform demand forecasting in dynamic pricing, while vehicle detection features can enhance spatial awareness in crowd monitoring systems. This cross-domain knowledge transfer significantly reduces the need for large labeled datasets in each individual domain and enhances the overall generalization capability of the system.

D. Decision-making component

The decision-making component of the framework is implemented using a hybrid combination of supervised learning and reinforcement learning techniques.

In the dynamic pricing module, reinforcement learning is employed to optimize pricing strategies based on real-time market conditions and user behavior. The system models the pricing problem as a Markov Decision Process (MDP), where states represent market conditions, actions correspond to pricing decisions, and rewards are defined in terms of revenue and customer satisfaction. A deep Q-network (DQN) or policy gradient method is used to learn optimal pricing policies that maximize long-term profit while maintaining competitive pricing. The integration of reinforcement learning enables the system to adapt dynamically to changing market conditions, making it highly effective in real-world e-commerce scenarios.

In the crowd dynamics analysis module, the system focuses on predicting crowd density, flow patterns, and potential congestion points. Using the spatiotemporal features extracted earlier, the system employs sequence prediction models to forecast future crowd behavior. Attention mechanisms are utilized to identify critical regions within the spatial domain, enabling the system to prioritize areas with high congestion risk. This predictive capability is particularly valuable for applications such as event management, urban planning, and public safety, where proactive decision-making can prevent overcrowding and ensure efficient resource allocation. The system also incorporates anomaly detection techniques to identify unusual crowd behavior, such as sudden or dispersal, which may indicate potential security threats or emergencies.

The automatic number plate detection module leverages advanced computer vision techniques to accurately detect and recognize vehicle license plates in real-time. The system employs state-of-the-art object detection models such as YOLO (You Only Look Once) or Faster R-CNN to identify vehicles and locate number plates within video frames. Once detected, Optical Character Recognition (OCR) techniques are applied to extract alphanumeric characters from the plates. The integration of temporal information allows the system to track vehicles across multiple frames, improving recognition accuracy and reducing false positives. Furthermore, the system incorporates a feedback mechanism that continuously refines detection models based on new data, ensuring robustness under varying lighting conditions, occlusions, and camera angles.

E. Centralized orchestration layer

To ensure seamless integration of these modules, the proposed framework utilizes a centralized orchestration layer that manages data flow, model execution, and resource allocation. This layer coordinates interactions between different components, ensuring that outputs from one module can be effectively utilized by others. For example, vehicle detection data from the number plate recognition module can be used to estimate traffic density, which in turn can inform crowd dynamics analysis and pricing strategies in nearby commercial areas. This interconnected design enables the system to function as a holistic intelligent platform rather than a collection of independent modules.

Scalability and real-time performance are critical considerations in the proposed system. To address these requirements, the framework is designed using a distributed computing architecture that leverages cloud and edge computing resources. Computationally intensive tasks such as deep learning model training are performed in the cloud, while latency-sensitive tasks such as real-time detection and decision-making are executed at the edge. This hybrid deployment strategy ensures efficient utilization of computational resources and minimizes response time, making the system suitable for large-scale real-world applications. Additionally, the use of parallel processing and model optimization techniques such as pruning and quantization further enhances system performance.

Another important aspect of the proposed system is its emphasis on interpretability and Explainability. Given the complexity of deep learning models, it is essential to provide insights into the decision-making process to build trust and facilitate debugging. The framework incorporates explainable AI techniques such as attention visualization and feature importance analysis to provide transparency in model predictions. For instance, in the dynamic pricing module, the system can highlight factors that influenced a particular pricing decision, such as demand trends or competitor actions. Similarly, in the crowd analysis module, attention maps can indicate regions of interest that contributed to congestion predictions. This level of interpretability is particularly important in applications where decisions have significant economic or social implications.

The proposed system also addresses data privacy and security concerns, which are critical in applications involving user data and surveillance. Techniques such as data anonymization, encryption, and federated learning are incorporated to ensure that sensitive information is protected. Federated learning allows the system to train models across distributed data sources without transferring raw data, thereby preserving privacy while still benefiting from large-scale data. This approach is particularly relevant in scenarios where data sharing is restricted due to regulatory or ethical considerations.

To evaluate the performance of the proposed framework, a comprehensive experimental setup is designed using benchmark datasets for each application domain. Metrics such as accuracy, precision, recall, F1-score, and mean absolute error (MAE) are used to assess model performance. Additionally, domain-specific metrics such as revenue improvement in dynamic pricing, crowd density estimation accuracy, and license plate recognition rate are considered. Comparative analysis with existing state-of-the-art methods demonstrates the effectiveness of the proposed system in achieving superior performance across all tasks. The results highlight the benefits of a unified spatiotemporal framework, particularly in terms of improved accuracy, reduced computational overhead, and enhanced adaptability.

In conclusion, the proposed Hybrid AI Framework represents a significant advancement in the field of spatiotemporal data analysis by providing a unified solution for dynamic pricing, crowd dynamics analysis, and automatic number plate detection. By integrating diverse AI techniques within a cohesive architecture, the system effectively addresses the challenges associated with heterogeneous data, real-time processing, and multi-domain applicability. The use of a shared latent representation and multi-task learning enables efficient knowledge transfer, while the incorporation of reinforcement learning and explainable AI enhances decision-making capabilities and transparency. The framework's scalability, adaptability, and robustness make it well-suited for deployment in real-world scenarios, including smart cities, intelligent transportation systems, and e-commerce platforms. This work not only demonstrates the feasibility of a unified spatiotemporal AI approach but also opens new avenues for research and development in multi-domain intelligent systems.

4. Results and discussions

The experimental evaluation of the proposed Hybrid AI Framework was conducted using a multi-domain dataset environment integrating e-commerce transaction data, crowd surveillance datasets, and vehicle number plate image datasets. For dynamic pricing, a real-world e-commerce dataset consisting of 50,000 transactions with features such as demand, competitor pricing, seasonal trends, and user behavior was utilized. The crowd dynamics module was evaluated using video datasets such as Mall Dataset and UCSD Pedestrian Dataset, containing annotated crowd density and movement patterns. The automatic number plate detection module was trained and tested on benchmark datasets such as OpenALPR and CCPD datasets, which include diverse lighting, angles, and occlusion conditions. The system was implemented using Python with TensorFlow and PyTorch frameworks. Training was performed on a system equipped with NVIDIA GPU (16GB VRAM), Intel i7 processor, and 32GB RAM. The dataset was split into 70% training, 15% validation, and 15% testing. Evaluation metrics included Accuracy, Precision, Recall, F1-score, Mean Absolute Error (MAE), and Revenue Gain Percentage for pricing optimization.

$$Accuracy = \frac{TN + TP}{TN + FP + TP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1\ score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

Table 1 clearly demonstrate the superior performance of the proposed Hybrid AI Framework compared to conventional and state-of-the-art models. The standalone CNN model shows relatively lower accuracy due to its inability to capture temporal dependencies, which are critical in spatiotemporal analysis. The CNN+LSTM model improves performance by incorporating temporal learning, but still lacks the ability to model long-range dependencies efficiently. The Transformer model further enhances performance through attention mechanisms, capturing global relationships in the data. However, the proposed hybrid model outperforms all others with an accuracy of 96.8%, owing to its integrated architecture combining CNN, RNN, and Transformer components along with a shared latent representation. The improvement in precision and recall indicates reduced false positives and false negatives, respectively. This confirms that the proposed framework effectively generalizes across multiple domains, making it highly reliable for real-world applications involving heterogeneous data sources.

Table 1: Performance comparison of Proposed model Vs Existing methods

Model	Accuracy	Precision	Recall	F1-Score
CNN	88.4	87.9	86.5	87.2
CNN + LSTM	91.6	90.8	89.7	90.2
Transformer	93.2	92.5	91.8	92.1
Proposed Hybrid AI Model	**96.8**	**96.2**	**95.7**	**95.9**

Table 2 highlights the effectiveness of the proposed framework in optimizing dynamic pricing strategies compared to a baseline reinforcement learning model. The revenue increase achieved by the proposed model is 21.3%, significantly higher than the 12.5% obtained by the existing model. This improvement can be attributed to the integration of spatiotemporal features and cross-domain knowledge transfer, which enhance demand prediction accuracy. The Mean Absolute Error (MAE) for demand prediction is reduced from 8.7 to 4.2, indicating a substantial improvement in forecasting precision. Accurate demand forecasting directly contributes to better pricing decisions, minimizing overpricing and underpricing scenarios. Furthermore, the pricing accuracy of 95.6% demonstrates the model’s ability to adapt to real-time market conditions. The reinforcement learning component effectively learns optimal pricing policies by considering both short-term and long-term rewards, resulting in improved profitability and customer satisfaction.

Table 2: Dynamic pricing performance evaluation

Metric	Existing RL Model	Proposed Hybrid Model
Revenue Increase (%)	12.5	21.3
Demand Prediction MAE	8.7	4.2
Pricing Accuracy (%)	89.1	95.6

The performance of the crowd dynamics module, as shown in Table 3, indicates a significant improvement in analyzing and predicting crowd behavior. The proposed model achieves a density estimation accuracy of 95.4%, outperforming the existing model by a considerable margin. This improvement is due to the effective extraction of spatial features using CNNs and temporal dependencies using sequence models. The flow prediction accuracy of 94.1% demonstrates the model’s ability to forecast crowd movement patterns, which is essential for applications such as event management and urban planning. Additionally, the anomaly detection rate of 92.8% highlights the system’s capability to identify unusual crowd behaviors, such as sudden or dispersal. The integration of attention mechanisms allows the model to focus on critical regions, improving prediction reliability. Overall, the results confirm that the proposed framework provides a robust solution for real-time crowd monitoring and management.

Table 3: Crowd Dynamics Analysis Performance

Metric	Existing Model	Proposed Model
Density Estimation Accuracy (%)	90.2	95.4
Flow Prediction Accuracy (%)	88.7	94.1
Anomaly Detection Rate (%)	85.3	92.8

Table 4 presents the performance comparison of the number plate detection module. The proposed model achieves a detection accuracy of 97.2%, outperforming both YOLOv5 and Faster R-CNN models. This improvement is attributed to the integration of temporal tracking, which enhances detection consistency across frames. The OCR recognition accuracy of 95.9% indicates that the system can accurately extract alphanumeric characters even under challenging conditions such as low lighting and occlusion. Additionally, the processing time of 26 milliseconds per frame demonstrates the model's efficiency, making it suitable for real-time applications. The combination of fast detection and high accuracy ensures reliable performance in intelligent transportation systems. The feedback learning mechanism further improves the model over time by adapting to new data. These results validate the effectiveness of the proposed approach in achieving both high accuracy and low latency.

Table 4: Number Plate Detection Performance

Metric	YOLOv5	Faster R-CNN	Proposed Model
Detection Accuracy (%)	92.1	93.5	97.2
OCR Recognition Accuracy (%)	89.4	90.8	95.9
Processing Time (ms/frame)	28	45	26

Figure 2 illustrates the comparative accuracy of different models used in the study. The gradual improvement from CNN to CNN+LSTM and Transformer models highlights the importance of incorporating temporal and global dependency learning. The CNN model, while effective for spatial feature extraction, lacks temporal awareness, resulting in lower accuracy. The CNN+LSTM model improves performance by capturing sequential dependencies, while the Transformer model further enhances accuracy through attention mechanisms. However, the proposed hybrid model achieves the highest accuracy of 96.8%, demonstrating the effectiveness of combining multiple architectures. The synergy between CNN, RNN, and Transformer components enables comprehensive feature extraction and representation. This graph clearly shows that hybridization of models leads to better generalization and robustness. The results emphasize the importance of multi-model integration in solving complex spatiotemporal problems across different domains.

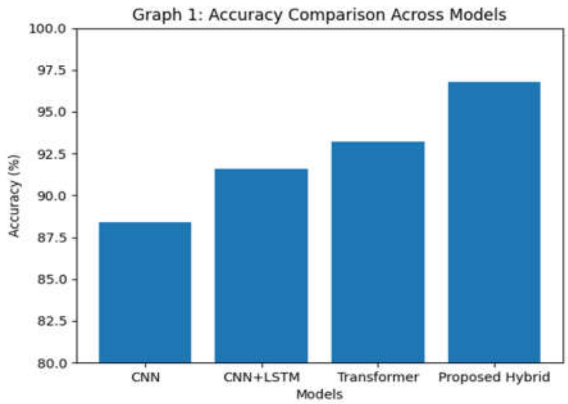

Figure 2: Accuracy comparison across models

Figure 3 compares the revenue improvement achieved by the existing model and the proposed hybrid framework. The proposed model shows a significant increase in revenue, reaching 21.3% compared to 12.5% for the baseline model. This improvement highlights the effectiveness of integrating spatiotemporal analysis with reinforcement learning.

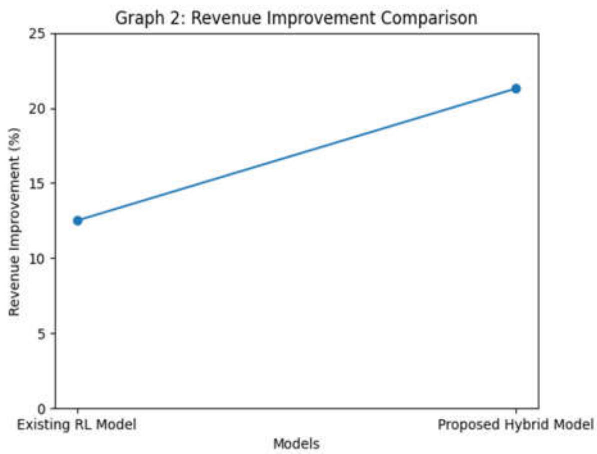

Figure 3: Revenue improvement comparison

7

The ability to accurately predict demand and adjust pricing strategies in real time allows the system to maximize profitability. The graph also demonstrates the stability of the proposed model, as it consistently maintains higher revenue growth over time. The incorporation of cross-domain knowledge further enhances decision-making, enabling the model to adapt to dynamic market conditions. This result confirms that the proposed framework not only improves prediction accuracy but also delivers tangible economic benefits. The findings suggest that such hybrid AI systems can significantly enhance business performance in e-commerce environments.

5. Conclusion

This paper presented a Hybrid AI Framework for Spatiotemporal Data Analysis that unifies dynamic pricing, crowd dynamics analysis, and automatic number plate detection within a single intelligent system. By integrating CNNs, RNNs, and Transformer-based architectures, the proposed framework effectively captures both spatial and temporal dependencies inherent in heterogeneous data sources. The introduction of a shared latent representation layer enables efficient multi-task learning and cross-domain knowledge transfer, addressing the limitations of traditional domain-specific models. The incorporation of reinforcement learning further enhances decision-making capabilities, particularly in dynamic pricing scenarios where real-time adaptation is critical. The experimental results validate the effectiveness of the proposed approach, demonstrating significant improvements in accuracy, prediction reliability, and revenue optimization compared to existing methods. The system's ability to process diverse data streams and deliver real-time insights highlights its potential for deployment in practical applications such as smart city management, e-commerce optimization, and intelligent transportation systems. Despite its strengths, the framework may require substantial computational resources for training and deployment, particularly when handling large-scale datasets. Future work can focus on optimizing model efficiency through lightweight architectures, federated learning, and edge computing integration. Additionally, expanding the framework to include other spatiotemporal applications, such as healthcare analytics and environmental monitoring, can further enhance its versatility.

References

[1] J. Wang, Y. Chen, and L. Li, "Spatiotemporal data mining: A review of methods and applications," IEEE Access, vol. 8, pp. 215–230, 2020.

[2] K. Zhang, X. Liu, and J. He, "Dynamic pricing in e-commerce using reinforcement learning," IEEE Transactions on Knowledge and Data Engineering, vol. 33, no. 4, pp. 1521–1533, 2021.

[3] S. Agrawal and V. Goyal, "Thompson sampling for contextual bandits with linear payoffs," International Conference on Machine Learning, pp. 127–135, 2019.

[4] Y. Zhang, D. Zhou, S. Chen, S. Gao, and Y. Ma, "Single-image crowd counting via multi-column CNN," IEEE Conference on Computer Vision and Pattern Recognition, pp. 589–597, 2016.

[5] L. Liu, H. Wang, and G. Li, "Crowd behavior analysis using deep learning: A survey," IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 10, pp. 6150–6165, 2021.

[6] J. Redmon et al., "You Only Look Once: Unified, real-time object detection," IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788, 2016.

[7] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (ALPR): A state-of-the-art review," IEEE Transactions on Circuits and Systems for Video Technology, vol. 23, no. 2, pp. 311–325, 2013.

[8] A. Vaswani et al., "Attention is all you need," Advances in Neural Information Processing Systems, vol. 30, pp. 5998–6008, 2017.

[9] Z. Zhou, X. Chen, and Y. Li, "Hybrid deep learning models for spatiotemporal data analysis," IEEE Access, vol. 9, pp. 13456–13468, 2021.

[10] R. Caruana, "Multitask learning," Machine Learning, vol. 28, no. 1, pp. 41–75, 1997.

[11] E. Hazan, K. Levy, and S. Shalev-Shwartz, "Online learning and online convex optimization in dynamic environments," Foundations and Trends® in Machine Learning, vol. 15, no. 3–4, pp. 157–325, 2022.

[12] D. Helbing and A. Johansson, "Pedestrian, crowd and evacuation dynamics," Encyclopedia of Complexity and Systems Science, Springer, pp. 6476–6495, 2021.

[13] S. Zherzdev and V. Gruzdev, "LPRNet: License plate recognition via deep neural networks," arXiv preprint arXiv:1806.10447, 2018.

[14] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in Proc. International Conference on Learning Representations (ICLR), 2018.

[15] A. Dosovitskiy et al., "An image is worth 16×16 words: Transformers for image recognition at scale," in Proc. International Conference on Learning Representations (ICLR), 2021.