

Article

QSA-QConvLSTM: A Quantum Computing-Based Approach for Spatiotemporal Sequence Prediction

Wenbin Yu, Zongyuan Chen, Chengjun Zhang and Yadang Chen

Special Issue

Quantum Information Processing and Machine Learning



Edited by

Dr. Wenbin Yu, Dr. Yadang Chen and Dr. Chengjun Zhang



Article

QSA-QConvLSTM: A Quantum Computing-Based Approach for Spatiotemporal Sequence Prediction

Wenbin Yu ^{1,*} , Zongyuan Chen ², Chengjun Zhang ³  and Yadang Chen ⁴ ¹ School of Software, Nanjing University of Information Science and Technology, Nanjing 210044, China² School of Computer Science, Nanjing University of Information Science and Technology, Nanjing 210044, China; 202212200029@nuist.edu.cn³ Wuxi Institute of Technology, Nanjing University of Information Science and Technology, Wuxi 214000, China; zcj@nuist.edu.cn⁴ Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAET), Nanjing University of Information Science and Technology, Nanjing 210044, China; adamchen@nuist.edu.cn

* Correspondence: ywb@nuist.edu.cn

Abstract: The ability to capture long-distance dependencies is critical for improving the prediction accuracy of spatiotemporal prediction models. Traditional ConvLSTM models face inherent limitations in this regard, along with the challenge of information decay, which negatively impacts prediction performance. To address these issues, this paper proposes a QSA-QConvLSTM model, which integrates quantum convolution circuits and quantum self-attention mechanisms. The quantum self-attention mechanism maps query, key, and value vectors using variational quantum circuits, effectively enhancing the ability to model long-distance dependencies in spatiotemporal data. Additionally, the use of quantum convolution circuits improves the extraction of spatial features. Experiments on the Moving MNIST dataset demonstrate the superiority of the QSA-QConvLSTM model over existing models, including ConvLSTM, TrajGRU, PredRNN, and PredRNN v2, with MSE and SSIM scores of 44.3 and 0.906, respectively. Ablation studies further verify the effectiveness and necessity of the quantum convolution circuits and quantum self-attention modules, providing an efficient and accurate approach to quantized modeling for spatiotemporal prediction tasks.



Academic Editor: Gabriel Luque

Received: 27 January 2025

Revised: 27 February 2025

Accepted: 4 March 2025

Published: 6 March 2025

Citation: Yu, W.; Chen, Z.; Zhang, C.; Chen, Y. QSA-QConvLSTM: A Quantum Computing-Based Approach for Spatiotemporal Sequence Prediction. *Information* 2025, 16, 206. <https://doi.org/10.3390/info16030206>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: quantum self-attention mechanism; quantum convolution circuits; spatiotemporal prediction; QSA-QConvLSTM

1. Introduction

Spatiotemporal sequence forecasting plays a crucial role in various fields such as meteorology, traffic management, and climate modeling. Particularly when dealing with complex spatiotemporal data, long-term dependencies and spatial patterns have a significant impact on prediction accuracy. Traditional models, such as ConvLSTM and its variant PredRNN, have achieved remarkable success in addressing these challenges by integrating convolutional layers with recurrent architectures. However, due to inherent structural limitations, these models often struggle to effectively capture long-range dependencies, leading to the gradual attenuation of information over time and adversely affecting their performance on long-sequence predictions [1,2]. To address these challenges, attention mechanisms have emerged in recent years as a promising solution, particularly for modeling long-range dependencies. By dynamically focusing on the most relevant information within a sequence, attention mechanisms effectively mitigate the limitations of

traditional methods in handling temporal dependencies [3,4]. Among these, self-attention mechanisms have proven particularly effective in capturing long-range dependencies in sequential data by establishing direct connections between elements across the entire sequence, thereby enhancing the flow of information [5]. Nevertheless, despite their advantages, classical self-attention mechanisms face significant computational challenges in practice. Specifically, their computational complexity increases quadratically with sequence length, limiting their scalability and applicability to large-scale datasets [6]. To overcome these limitations, this study proposes a novel quantum self-attention (QSA) mechanism that leverages the parallelism and superposition properties of quantum computing to enhance the modeling capacity while reducing computational complexity. The QSA mechanism employs variational quantum circuits (VQCs) to replace the classical mapping operations in the query, key, and value vectors, thereby improving the modeling of dependencies and enhancing the extraction of spatial features. Compared to traditional methods, the QSA mechanism benefits from quantum computing to capture more complex spatiotemporal dependencies, significantly improving both model accuracy and computational efficiency [7]. Building upon this innovative mechanism, we further propose the QSA-QConvLSTM model, which integrates the QSA mechanism with the quantum convolutional circuit (QConv) introduced in prior studies to address the complexity of spatiotemporal data. By simulating convolutional operations and incorporating the advantages of quantum computing, QConv effectively extracts spatial features with greater efficiency, enhancing overall model performance [8,9]. This hybrid model not only improves the predictive accuracy for spatiotemporal forecasting tasks but also significantly reduces computational resource consumption, thereby improving computational efficiency.

2. Related Work

2.1. Spatiotemporal Sequence Forecasting

Recurrent Neural Networks (RNNs) serve as foundational models for sequence data processing. By recursively passing the previous state to the current time step, RNNs can effectively capture temporal dependencies [10]. However, RNNs are prone to issues such as vanishing or exploding gradients when handling long sequences, which limits their performance in modeling long-range dependencies [11]. Long Short-Term Memory networks (LSTMs), an extension of RNNs, address the vanishing gradient problem by introducing gating mechanisms (such as input, forget, and output gates), enabling the model to capture long-term dependencies more effectively [12]. Despite significant improvements over RNNs, LSTMs still face challenges in terms of computational complexity and their ability to model long sequences. Gated Recurrent Units (GRUs) represent a simplified version of LSTMs, with fewer parameters and a higher computational efficiency. In certain tasks, GRUs have demonstrated performance comparable to or even better than that of LSTMs [13]. However, this simplification means that GRUs may be less capable in handling complex temporal dependencies. To further enhance the performance of spatiotemporal sequence forecasting, Convolutional LSTM (ConvLSTM) combines Convolutional Neural Networks (CNNs) with LSTMs, allowing for the simultaneous capture of spatial features and temporal dependencies. This architecture is particularly suitable for tasks such as video prediction and weather data modeling [1]. Despite the success of ConvLSTM in spatiotemporal data modeling, it still faces high computational costs when dealing with long time sequences and struggles with non-uniform spatiotemporal dependencies. TrajGRU, designed specifically for trajectory prediction tasks, adopts the GRU structure and further optimizes spatial information modeling. It effectively predicts the trajectories of moving objects but is limited in modeling complex spatiotemporal dynamics [14]. In contrast, Pre-dRNN introduces spatiotemporal convolutional LSTMs based on ConvLSTM, utilizing this

architecture to capture more complex spatiotemporal dependencies in video sequences and achieve better performance in tasks like video prediction [15]. However, PredRNN suffers from high computational costs, especially for long-term forecasting tasks, where the computational resource consumption becomes excessive. Building upon PredRNN, PredRNNv2 introduces a dual-stream LSTM architecture, enabling the model to simultaneously handle both local and global spatiotemporal dependencies, thereby improving forecasting accuracy. This approach has demonstrated superior performance across various tasks, but its more complex network structure brings higher computational overhead [16]. Consequently, when processing large-scale datasets, optimizing and managing computational resources remain critical considerations.

2.2. Quantum Computing in Machine Learning

With the rapid development of quantum computing, its integration with classical computing models has become a hot research topic. Quantum computing not only provides more efficient solutions to specific problems but also has the potential to improve the performance of classical machine learning and optimization models to some extent. On the one hand, quantum computing can accelerate the training process of classical models through quantum algorithms; on the other hand, quantum computing itself offers new perspectives for classical models, driving the emergence of hybrid quantum–classical algorithms. In the field of machine learning, quantum computing has been widely used to enhance the efficiency of classical algorithms. For instance, Quantum Support Vector Machines (QSVMs) can provide better classification performance when handling high-dimensional data [17], with the quantum advantage primarily stemming from the kernel function computation in high-dimensional Hilbert spaces. Compared to classical Support Vector Machines (SVMs), QSVMs can handle larger datasets and extract higher-dimensional features through quantum computation, resulting in better performance in certain tasks than classical methods. Similarly, Quantum Reinforcement Learning (QRL) combines the parallelism of quantum computing with the strategy optimization framework of classical reinforcement learning, utilizing quantum computing to accelerate the search and optimization processes of strategies [18]. Quantum Annealing and Quantum Approximate Optimization Algorithms (QAOAs) offer novel solutions to classical optimization problems. QAOAs leverage the superposition and entanglement properties of quantum bits to provide greater potential in solving combinatorial optimization problems compared to classical algorithms. For example, when solving NP-hard problems like the Traveling Salesman Problem (TSP) and Maximum Cut Problem, QAOAs can effectively provide closer approximations to the optimal solution, outperforming classical heuristic algorithms in some specific instances [7]. Moreover, Quantum Annealing, by simulating the evolution of quantum systems in low-energy states, offers a distinct optimization approach to classical annealing algorithms, especially suited for large-scale optimization problems. Quantum computing has also shown promise in improving classical neural network models. Quantum Convolutional Neural Networks (QCNNs) integrate quantum computing with Convolutional Neural Networks (CNNs), taking full advantage of quantum computers' ability to process high-dimensional spaces when dealing with spatiotemporal data with high-dimensional features. Research indicates that QCNNs can enhance the learning and reasoning efficiency of models in tasks such as image classification and video analysis by utilizing the superposition and entanglement effects of quantum bits [8,19]. This innovation not only provides a new computational architecture for classical neural networks but also offers a feasible path for integrating quantum computing into practical applications. In summary, quantum computing, by accelerating the computational process of classical algorithms, improving optimization methods, and offering new computational perspectives

for classical models, has driven the development of hybrid quantum–classical algorithms. In the future, as quantum hardware continues to advance, the deep integration of quantum computing with classical computing models will bring more significant performance improvements across various fields.

3. Method

Although classical attention mechanisms excel in capturing long-range dependencies and modeling complex data interactions, their computational complexity grows exponentially as the input data dimensions increase. This becomes particularly challenging when handling high-dimensional or large-scale data, as the memory and computational resource requirements are extremely demanding. Quantum computing, with its unique properties of superposition and parallelism, can offer significant advantages in processing high-dimensional data. Variational quantum circuits (VQCs) efficiently model complex distributions through a compact parameterized approach, reducing computational complexity and enabling superior feature extraction capabilities within smaller memory spaces. At the same time, quantum entanglement enhances the ability to model inter-data correlations, allowing the quantum attention mechanism to capture global dependencies more efficiently and address the shortcomings of classical attention mechanisms when dealing with high-dimensional, complex data [20,21].

3.1. Quantum Self-Attention Mechanism Process

This paper proposes a quantum self-attention mechanism based on the integration of quantum and classical computing, enhancing model performance and practical feasibility by quantumizing certain modules of the classical self-attention mechanism. In classical self-attention, query, key, and value vectors are generated through mapping operations. In the quantum self-attention mechanism, variational quantum circuits (VQCs) replace traditional linear transformations or convolution operations, realizing the quantumized mapping process. The computational process of this mechanism is shown in Figure 1. First, the classical input is encoded into a quantum state. Then, a variational quantum circuit is used to perform vector mapping. Subsequently, quantum measurements are employed to obtain similarity, and attention scores are computed using the cross-covariance matrix. After normalizing the scores, the attention weights are obtained. Finally, the weights are applied to the value vectors, and a weighted sum is taken to produce the final result.

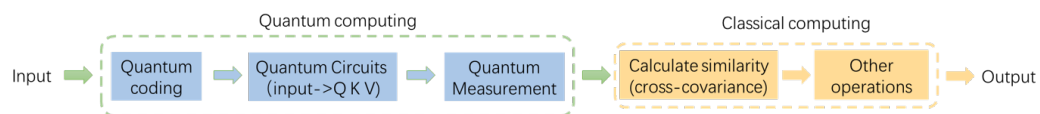


Figure 1. Flowchart of quantum self-attention mechanism.

In the self-attention mechanism, the commonly used method for calculating similarity is dot product, but this significantly increases both time and space complexity. For spatiotemporal data, using the cross-covariance matrix to calculate similarity is more efficient. This not only reduces model complexity but also simplifies the training process while maintaining prediction accuracy. Cross-covariance is a tool for measuring the degree of association between different variables. A positive value indicates a positive correlation, and a negative value indicates a negative correlation; additionally, the larger the absolute value, the stronger the association. A value of zero indicates no correlation. Unlike covariance, which is used for a univariate analysis, cross-covariance is used for analyzing relationships between multiple variables and is widely applied in statistics and signal processing fields. For instance, a cross-covariance matrix can effectively measure the similarity between signals and can be used to infer unknown characteristics from known signals. Therefore,

to calculate similarity using a cross-covariance matrix, the computational formula shown in Equations (1) and (2) is used. In this equation, \hat{K} and \hat{Q} represent the normalized key and query matrices, respectively. This normalization helps to enhance the stability of the training process. Additionally, τ is the scaling factor, which assists in distributing the attention weights more evenly.

The attention mechanism is computed as

$$\text{Attention}(Q, K, V) = V \cdot \mathcal{A}(K, Q), \tag{1}$$

where the function $\mathcal{A}(K, Q)$ is defined as the softmax of the scaled dot product between \hat{K} and \hat{Q} :

$$\mathcal{A}(K, Q) = \text{Softmax}\left(\frac{\hat{K}^\top \hat{Q}}{\tau}\right). \tag{2}$$

3.2. Quantumization of Linear Mapping

In the self-attention mechanism, mapping from the input to the query, key, and value vectors is typically achieved through linear transformations, where the input data are projected into different subspaces using weight matrices. This method is widely used in fields such as natural language processing (as shown on the left side of Figure 2). To enhance feature representation and reduce computational complexity, this study introduces variational quantum circuits (VQCs) to replace traditional linear transformations, completing the quantumized mapping process. The specific structure of this quantumized mapping is shown on the right side of Figure 2.

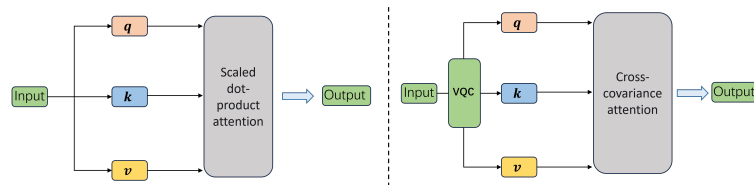


Figure 2. Quantization of linear mapping.

Figure 3 illustrates the variational quantum circuit (VQC) used in this study. The circuit consists of quantum rotation gates and CNOT gates. The rotation gates are chosen as the Y-gate and Z-gate, which can alter the amplitude and phase of the corresponding quantum bits. The CNOT gate is a controlled gate, consisting of a control qubit and a target qubit. It changes the state of the target qubit based on the state of the control qubit. By placing CNOT gates in the circuit, quantum entanglement between quantum lines can be achieved. The circuit blocks within the dashed frame can be repeated multiple times to increase the depth and expressive power of the circuit.

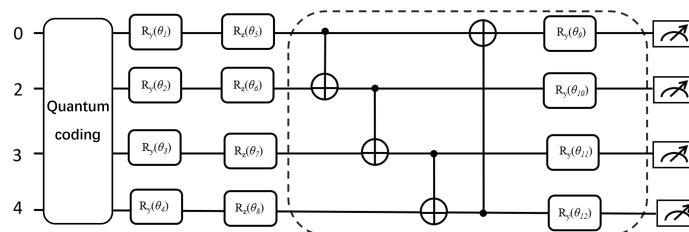


Figure 3. Variational quantum circuit diagram.

3.3. Quantumization of Convolutional Mapping

Mapping from the input to the corresponding vectors is achieved through quantum convolution operations, which are particularly suitable for data like images and videos.

As shown in Figure 4, the quantum convolution circuit replaces classical convolution operations with a quantized alternative through a hierarchical structure:

1. Encoding Layer: Converts classical data into quantum state representations.
2. Convolution Kernel Layer: Uses parameterized quantum gates to perform feature mapping.
3. Dynamic Adjustment Layer: Adapts the number of quantum bits and encoding schemes based on task requirements.

The output from the quantum convolutional layer passes through a softmax function, which normalizes the result. The softmax function is typically used to transform the values into a probability distribution, ensuring that the output vectors are weighted appropriately for the final prediction. This step is essential for computing the attention scores in the context of quantum self-attention mechanisms, allowing the model to focus on the most relevant features from the input data. This hierarchical design retains the spatial locality features of classical convolutions while utilizing the parallel computation advantages offered by quantum superposition in a high-dimensional Hilbert space. Furthermore, the use of softmax enhances the model’s ability to dynamically adjust and prioritize different parts of the input data for improved performance.

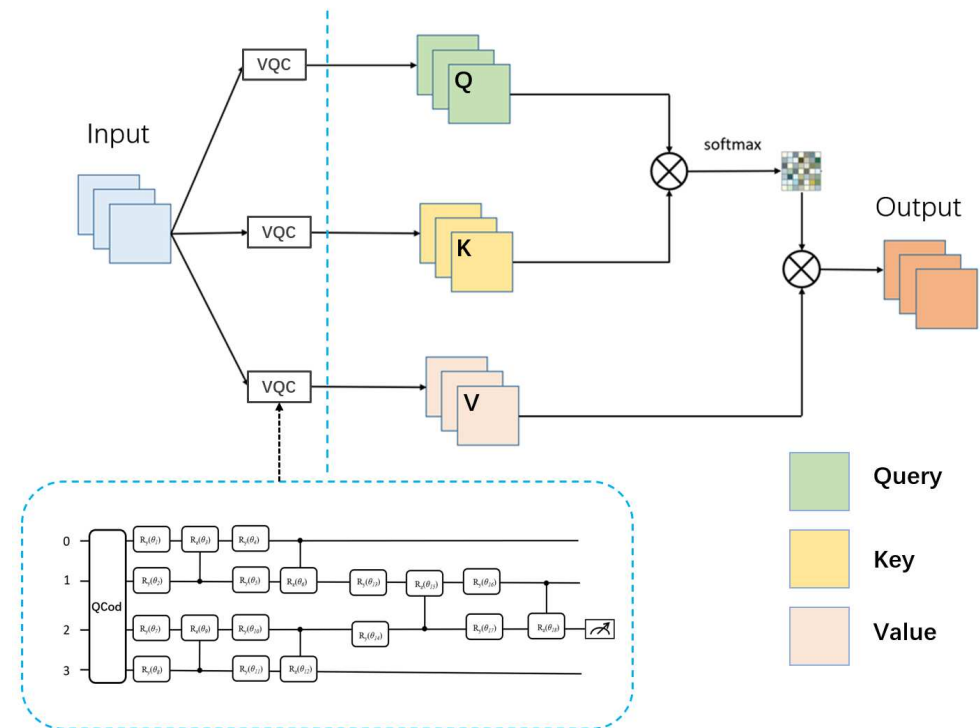


Figure 4. Quantization of convolution map.

The numbers 0, 1, 2, 3 represent qubit indices, while $\theta_1, \theta_2, \dots$ are trainable parameters of quantum gates (such as R_y and R_z) used to adjust quantum states. Through these rotation and controlled gates, the circuit performs feature mapping and quantum computation to enhance data representation.

3.4. Model Methodology

The essence of the ConvLSTM model lies in its recurrently connected structure. This architecture makes it difficult for the model to effectively capture long-range dependencies, thereby limiting its analytical and predictive capabilities. In contrast, the attention mechanism processes information based on weights, independent of temporal sequence lengths, and it can better model relationships between data points. To address this limitation, this

study introduces a quantized adaptation of the classical self-attention mechanism. By replacing the classical mapping operations with variational quantum circuits (VQCs), we integrate the quantum self-attention mechanism into the ConvLSTM model and propose the QSA-ConvLSTM model. This approach enhances the model’s ability to capture long-range dependencies, thereby improving both its prediction accuracy and computational efficiency. The internal structure of the QSA-ConvLSTM model is illustrated in Figure 5. The QSA module is a quantum self-attention module that performs quantized processing on convolutional mappings, where the VQC corresponds to the quantum circuit depicted in Figure 4.

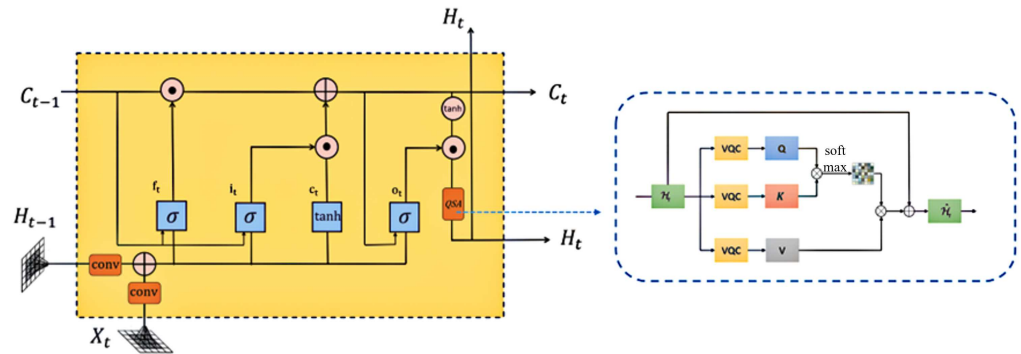


Figure 5. QSA-ConvLSTM model.

The internal computational methods for each state in the QSA-ConvLSTM model are described in Equations (3)–(10). Here, QSA refers to the quantum self-attention module, and \mathcal{X}_t and \mathcal{H}_{t-1} represent the features aggregated by the QSA module. The computational steps are as follows:

$$\hat{\mathcal{X}}_t = \text{QSA}(\mathcal{X}_t), \tag{3}$$

$$\hat{\mathcal{H}}_{t-1} = \text{QSA}(\mathcal{H}_{t-1}), \tag{4}$$

$$i_t = \sigma(W_{xi} * \hat{\mathcal{X}}_t + W_{hi} * \hat{\mathcal{H}}_{t-1} + b_i), \tag{5}$$

$$f_t = \sigma(W_{xf} * \hat{\mathcal{X}}_t + W_{hf} * \hat{\mathcal{H}}_{t-1} + b_f), \tag{6}$$

$$g_t = \tanh(W_{xc} * \hat{\mathcal{X}}_t + W_{hc} * \hat{\mathcal{H}}_{t-1} + b_c), \tag{7}$$

$$C_t = f_t \circ C_{t-1} + i_t \circ g_t, \tag{8}$$

$$o_t = \sigma(W_{xo} * \hat{\mathcal{X}}_t + W_{ho} * \hat{\mathcal{H}}_{t-1} + b_o), \tag{9}$$

$$\mathcal{H}_t = o_t \circ \tanh(C_t). \tag{10}$$

This paper proposes the QSA-QConvLSTM model, which leverages quantum convolutional circuits to enhance the model’s ability to extract spatial features and employs the quantum self-attention mechanism to better capture long-range dependencies. This approach improves the accuracy of prediction tasks.

3.5. Quantum Circuit Design

In the design of quantum convolutional circuits, we selected a quantum circuit with four qubits. To further justify this choice over other possible qubit numbers, we conducted a comparative experiment. In this experiment, based on a hierarchical circuit structure, we constructed quantum convolutional circuits with different configurations, namely, a single-layer two-qubit circuit, a two-layer four-qubit circuit, and a three-layer six-qubit circuit, as illustrated in Figure 6. Ultimately, by comparing the performance of circuits with different hierarchical structures, we identified the optimal circuit configuration.

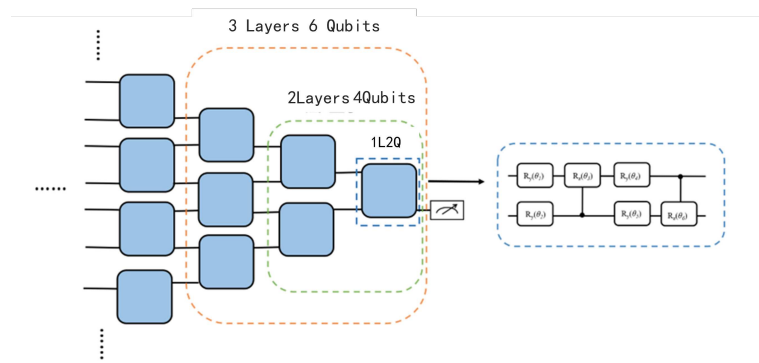


Figure 6. Multi-layer quantum circuit.

Table 1 summarizes the experimental results. The findings indicate that the single-layer two-qubit circuit exhibits the lowest accuracy. This is because a reduced number of qubits leads to significant information loss, thereby impairing the model’s learning and prediction capabilities. Although the three-layer six-qubit circuit achieves similar predictive performance to the two-layer four-qubit circuit, its depth and the number of parameters increase substantially, resulting in a significant rise in quantum resource consumption. Furthermore, for deep neural network models, excessively complex quantum circuits may not be suitable. Therefore, considering both prediction accuracy and resource efficiency, we ultimately selected the two-layer four-qubit circuit configuration.

Table 1. Comparison of experimental results with different quantum circuit configurations.

Circuit Structure	MSE	SSIM
1 Layer, 2 Qubits	96.8	0.715
2 Layers, 4 Qubits	95.6	0.718
3 Layers, 6 Qubits	95.9	0.717

4. Experiment

4.1. Moving MNIST Dataset

In this study, the Moving MNIST dataset is used for experiments. The dataset contains 10,000 training sequences and 2000 testing sequences. Each sequence consists of 20 frames of 64×64 -pixel images, featuring two digits that move at a constant speed on a black background and bounce off the boundaries. The digits may overlap, adding to the prediction difficulty. The first 10 frames of each sequence are used as input to predict the subsequent 10 frames, and all pixel values are normalized to the range [0, 1]. The complex spatiotemporal dynamics of this dataset provide a solid benchmark for evaluating model performance.

4.2. Experimental Results

The experiments in this study were conducted on the Moving MNIST dataset. The Mean Squared Error (MSE) and Structural Similarity Index Measure (SSIM) were used as evaluation metrics to compare the performance of the QSA-QConvLSTM model against that of four other models: ConvLSTM, TrajGRU, PredRNN, and PredRNN v2. The experimental results are presented in Table 2, where the MSE and SSIM values represent the averages over the predicted 10 frames. From the results, it is evident that the QSA-QConvLSTM model outperforms the other four models. Compared to the classic ConvLSTM model, QSA-QConvLSTM demonstrates a significant improvement in prediction accuracy.

Table 2. Comparison of experimental results.

Model	MSE	SSIM
ConvLSTM	101.1	0.705
TrajGRU	96.7	0.715
PredRNN	61.6	0.878
PredRNN v2	51.9	0.891
QSA-QConvLSTM	44.3	0.906

Figure 7 illustrates the trends of the evaluation metrics over 10 time steps. In the figure, (a) represents the MSE metric, while (b) represents the SSIM metric. As the time steps progress, the prediction performance of all models gradually deteriorates, indicated by an increasing trend in the MSE curve and a decreasing trend in the SSIM curve. However, the QSA-QConvLSTM model exhibits a clear advantage over the other models, as its MSE metric increases at a slower rate, and its SSIM metric decreases at a slower rate.

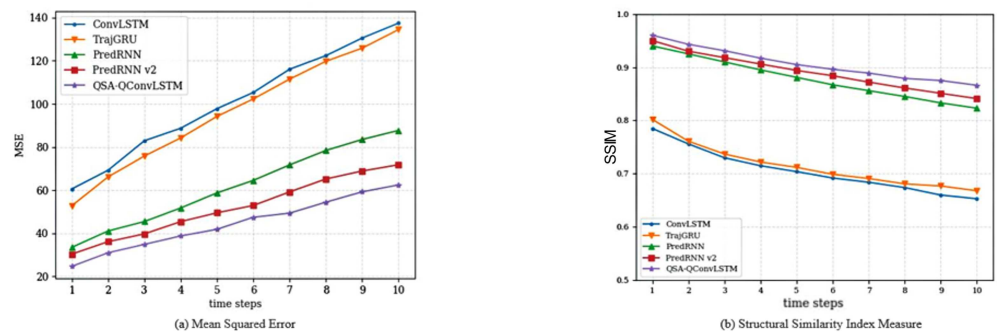


Figure 7. Curves of evaluation metric variations.

To provide a more intuitive demonstration of the prediction performance of the different models, two samples from the Moving MNIST dataset were selected for a visualization analysis. Figure 8 illustrates two prediction examples. In the figure, Input represents the first 10 input frames, and Ground Truth represents the actual subsequent 10 frames. In Figure 8, it is evident that the QSA-QConvLSTM model predicts the shapes and evolution patterns of digits more accurately than the other models. For instance, in Figure 8b, during the last five frames of the input sequence, the PredRNN and PredRNN v2 models fail to recognize the digit “8”, mistakenly identifying it as the digit “5”. By contrast, the QSA-QConvLSTM model successfully separates the digits and achieves a higher prediction accuracy.

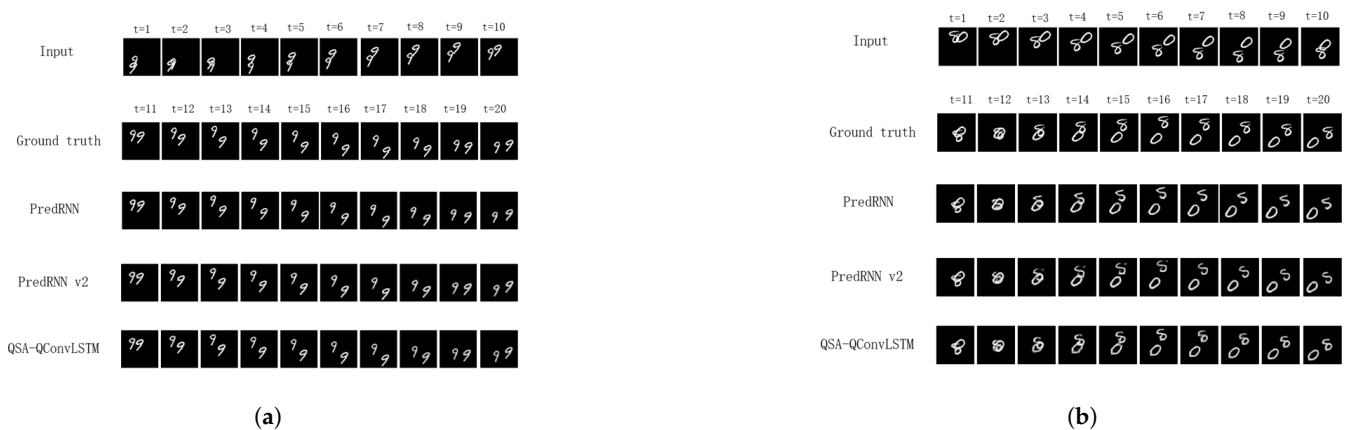


Figure 8. Display of prediction samples: (a) prediction example 1, (b) prediction example 2.

Ablation experiments were conducted on the Moving MNIST dataset to verify the effectiveness and necessity of the quantum convolutional circuit (QConv) and the quantum self-attention module (QSA). The ablation experiments were performed using a step-by-step addition and variable control approach. The results of the experiments are shown in Table 3, where “w” indicates the presence of the structure, and “w/o” indicates its absence. QConv represents the quantum convolutional circuit, and QSA represents the quantum self-attention module. From the results of the ablation experiments, it is evident that both the quantum convolutional circuit and the quantum self-attention module are effective. Applying these two methods to the ConvLSTM model significantly improves prediction accuracy.

Table 3. Ablation experiment results. Comparison of models with different configurations.

Method	MSE	SSIM
ConvLSTM	101.1	0.705
w QConv, w/o QSA	95.6	0.718
w/o QConv, w QSA	46.2	0.897
w QConv, w QSA	44.3	0.906

4.3. ERA5 Dataset Experiment and Computational Complexity Analysis

This study uses the ERA5 dataset [22] to validate the quantum acceleration effect of the QSA-QConvLSTM model. The experiment utilizes meteorological variables such as temperature, humidity, and wind speed, using the first 10 h of data to predict the next 10 h, and it compares the performance with that of baseline models such as ConvLSTM. The experimental results are summarized in Table 4, which show that combining QConv and QSA can significantly improve model performance. Specifically, the QSA-QConvLSTM model achieves the lowest MSE (74.5) and the highest SSIM (0.786), which is significantly improved compared to the baseline ConvLSTM (MSE: 98.3, SSIM: 0.730), demonstrating the effectiveness of the quantum module in spatiotemporal prediction tasks.

Table 4. Experimental results of QSA-QConvLSTM on ERA5 dataset.

Model	MSE	SSIM
ConvLSTM	98.3	0.730
w QConv, w/o QSA	87.9	0.753
w/o QConv, w QSA	83.2	0.775
w QConv, w QSA	74.5	0.786

In terms of computational complexity, QSA-QConvLSTM significantly reduces computational overhead through the quantum self-attention mechanism (QSA) and quantum convolutional module (QConv). The QSA mechanism leverages quantum parallelism, reducing the classical complexity of self-attention from $O(N^2d)$ to $O(Nd \log d)$, enabling a 20× computation speedup for long sequences (e.g., $N = 1000$). Similarly, QConv reduces the computational cost of classical convolution from $O(M^2m^2)$ to $O((\log_2 M)^4)$ by using hierarchical parameterized quantum circuits, achieving a 44% reduction in parameters compared to deeper configurations. In conclusion, QSA-QConvLSTM provides a high-performance solution for spatiotemporal sequence modeling by improving prediction accuracy while effectively reducing computational complexity, demonstrating the potential of quantum computing for large-scale data modeling.

5. Discussion

Deploying quantum models on Noisy Intermediate-Scale Quantum (NISQ) devices is both promising and challenging. Similar to the QSA-QConvLSTM model, the SA-DQAS framework [7,23] introduces the self-attention mechanism into quantum architecture search, highlighting the feasibility of deploying quantum models on NISQ devices. It enhances the stability and noise resilience of quantum circuits through dynamic architecture optimization, addressing one of the key challenges of NISQ hardware—quantum noise. The QSA-QConvLSTM model similarly benefits from quantum self-attention, optimizing the computation process through quantum parallelism, thereby reducing computational complexity while improving accuracy. This aligns with the findings of QAS [24], where quantum-enhanced models demonstrated better noise resilience and efficiency in NISQ environments. Despite the promising outlook, quantum noise and the limited number of qubits remain significant challenges. The QSA-QConvLSTM model reduces qubit dependency through hybrid classical–quantum algorithms, outperforming classical models such as ConvLSTM and PredRNN in prediction accuracy and computational efficiency. Quantum self-attention and convolution circuits offer a feasible approach for improving spatiotemporal sequence prediction on NISQ devices. Further research should focus on optimizing quantum circuit designs to address noise and scalability issues, ensuring the deployment of these models in real-world applications.

6. Conclusions

This study introduces the QSA-QConvLSTM model, which significantly improves prediction performance and computational efficiency in spatiotemporal sequence forecasting tasks by incorporating the quantum self-attention (QSA) mechanism and quantum convolutional circuit (QConv). The experimental results demonstrate that the QSA-QConvLSTM model outperforms traditional models, such as ConvLSTM and PredRNN, in terms of the MSE and SSIM metrics, validating the effectiveness of quantum modules in capturing long-range dependencies and spatial features. The QSA mechanism enhances the model's ability to capture complex dependencies in long time-series data by leveraging the parallelism and superposition properties of quantum computing, thereby avoiding the computational bottleneck faced by classical self-attention mechanisms when processing long sequences. In contrast, the QConv circuit simulates quantum convolution operations to further improve the efficiency of spatial feature extraction. The combination of these two modules enables the QSA-QConvLSTM model to achieve superior performance in spatiotemporal prediction tasks. However, despite the significant performance gains brought by quantum modules, the issue of quantum noise remains a challenge in the field of quantum computing. In the NISQ (Noisy Intermediate-Scale Quantum) era, quantum hardware limitations may introduce quantum noise, which can affect the stability and robustness of the model. Future research could focus on optimizing quantum circuit designs, improving the noise resilience of models, and exploring hybrid quantum–classical algorithms to address the limitations of current quantum computing hardware. In conclusion, the QSA-QConvLSTM model provides a novel approach for applying quantum computing to spatiotemporal sequence prediction and demonstrates the potential of quantum computing to improve model accuracy and efficiency. With the continuous advancement of quantum technology, quantum computing is expected to play an increasingly important role in complex spatiotemporal data modeling in the future.

Author Contributions: Conceptualization, W.Y. and Z.C.; methodology, Z.C.; validation, W.Y., Z.C. and Y.C.; investigation, Y.C.; resources, Z.C.; data curation, C.Z.; writing—original draft preparation, W.Y.; writing—review and editing, Z.C. and C.Z.; visualization, Y.C.; supervision, W.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of China, grant numbers 62473201; the Basic Research Program of Jiangsu, grant number BK20231142; and the Innovation Program for Quantum Science and Technology, grant number 2021ZD0302901.

Data Availability Statement: Data is contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

QSA	quantum self-attention
QConvLSTM	quantum convolutional LSTM
QCod	quantum coding

References

- Shi, X.; Chen, Z.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 802–810. [[CrossRef](#)]
- Ma, Y.; Tresp, V.; Zhao, L.; Wang, Y. Variational quantum circuit model for knowledge graph embedding. *Adv. Quantum Technol.* **2019**, *2*, 1800078. [[CrossRef](#)]
- Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**. [[CrossRef](#)]
- Ma, Y.; Tresp, V. Quantum machine learning algorithm for knowledge graphs. *ACM Trans. Quantum Comput.* **2021**, *2*, 1–28. [[CrossRef](#)]
- Lin, Z.; Feng, M.; Santos, C.N.D.; Yu, M.; Xiang, B.; Zhou, B.; Bengio, Y. A structured self-attentive sentence embedding. *arXiv* **2017**, arXiv:1703.03130.
- Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
- Farhi, E.; Goldstone, J.; Gutmann, S. A quantum approximate optimization algorithm. *arXiv* **2014**, arXiv:1411.4028.
- Cong, I.; Choi, S.; Lukin, M.D. Quantum convolutional neural networks. *Nat. Phys.* **2019**, *15*, 1273–1278. [[CrossRef](#)]
- Giovagnoli, A.; Tresp, V.; Ma, Y.; Schubert, M. Qneat: Natural evolution of variational quantum circuit architecture. In Proceedings of the Companion Conference on Genetic and Evolutionary Computation, Lisbon, Portugal, 15–19 July 2023; pp. 647–650.
- Medsker, L.R.; Jain, L. Recurrent Neural Networks. *Des. Appl.* **2001**, *5*, 64–67.
- Graves, A.; Graves, A. Long short-term memory. In *Supervised Sequence Labelling with Recurrent Neural Networks*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 37–45.
- Graves, A. Generating sequences with recurrent neural networks. *arXiv* **2013**, arXiv:1308.0850.
- Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
- Shi, X.; Gao, Z.; Lausen, L.; Wang, H.; Yeung, D.Y.; Wong, W.K.; Woo, W.C. Deep learning for precipitation nowcasting: A benchmark and a new model. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [[CrossRef](#)]
- Wang, Y.; Long, M.; Wang, J.; Gao, Z.; Yu, P.S. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal LSTMs. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 879–888.
- Wang, Y.; Long, M.; Wang, J.; Gao, Z.; Yu, P.S. Predrnn: A recurrent neural network for spatiotemporal predictive learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 2208–2225. [[CrossRef](#)] [[PubMed](#)]
- Rebentrost, P.; Mohseni, M.; Lloyd, S. Quantum support vector machine for big data classification. *Phys. Rev. Lett.* **2014**, *113*, 130503. [[CrossRef](#)] [[PubMed](#)]
- Dong, D.; Chen, C.; Li, H.; Tarn, T.J. Quantum reinforcement learning. *IEEE Trans. Syst. Man, Cybern. Part B (Cybernetics)* **2008**, *38*, 1207–1220. [[CrossRef](#)]
- Wei, S.; Chen, Y.; Zhou, Z.; Long, G. A quantum convolutional neural network on NISQ devices. *AAPPS Bull.* **2022**, *32*, 1–11. [[CrossRef](#)]
- Li, G.; Zhao, X.; Wang, X. Quantum self-attention neural networks for text classification. *Sci. China Inf. Sci.* **2024**, *67*, 142501. [[CrossRef](#)]
- Shi, S.; Wang, Z.; Li, J.; Li, Y.; Shang, R.; Zheng, H.; Zhong, G.; Gu, Y. A natural NISQ model of quantum self-attention mechanism. *arXiv* **2023**, arXiv:2305.15680.

22. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [[CrossRef](#)]
23. Sun, Y.; Wu, Z.; Ma, Y.; Tresp, V. Quantum architecture search with unsupervised representation learning. *arXiv* **2024**, arXiv:2401.11576.
24. Sun, Y.; Liu, J.; Wu, Z.; Ding, Z.; Ma, Y.; Seidl, T.; Tresp, V. SA-DQAS: Self-attention enhanced differentiable quantum architecture search. *arXiv* **2024**, arXiv:2406.08882.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.