Journal of Organizational and End User Computing
Volume 37 • Issue 1 • January-December 2025

Artificial Intelligence-Driven Optimization of Carbon Neutrality Strategies in Population Studies:

Employing Enhanced Neural Network Models With Attention Mechanisms

Sida Guo
https://orcid.org/0009-0008-7308-7879
Research Institute of Social Development, Southwestern University of Finance and Economics, China

Ziqi Zhong
https://orcid.org/0000-0002-3919-9999
London School of Economics and Political Science, UK

ABSTRACT

With the growing severity of global climate change, achieving carbon neutrality has become a central focus worldwide. The intersection of population studies and carbon neutrality introduces significant challenges in predicting and optimizing energy consumption, as demographic factors play a crucial role in shaping carbon emissions. This paper proposes a model based on a Region-based Convolutional Neural Network (RCNN) and Generative Adversarial Network (GAN), enhanced with a dual-stage attention mechanism for optimization. The model automatically extracts key features from complex demographic and carbon emission data, leveraging the attention mechanism to assign appropriate weights, thereby capturing the behavioral patterns and trends in energy consumption driven by population dynamics more effectively. By integrating multi-source data, including historical carbon emissions, population density, demographic trends, meteorological data, and economic indicators, experimental results demonstrate the model's outstanding performance across multiple datasets.

KEYWORDS

Artificial Intelligence, Deep Learning, Carbon Neutral, Fusion model, Two-Stage Attention Optimization, Data Analysis

INTRODUCTION

As global climate change intensifies, carbon neutrality has become an urgent goal for societies worldwide. Population growth, urbanization, and demographic shifts play a crucial role in shaping energy consumption patterns and carbon emissions. To mitigate climate change and reduce greenhouse gas emissions, various industries are actively seeking innovative solutions to minimize their environmental impact.

Deep learning, a critical branch of artificial intelligence, has made breakthrough achievements in image recognition, natural language processing, and speech recognition by mimicking the human nervous system's learning process from large datasets (Lyu et al., 2024; Shi et al., 2022; Xi et al.,

DOI: 10.4018/JOEUC.370801

2024). In the realm of environmental protection and sustainable development, deep learning is widely applied in energy management, climate simulation, and intelligent transportation, providing powerful tools and methodologies to address complex environmental challenges.

The sports industry involves significant energy consumption and carbon emissions during the operation of events, activities, and facilities. Therefore, optimizing energy usage and reducing carbon footprints are crucial in the carbon neutrality process (Karpov et al., 2019). In recent years, deep learning technology has been gradually introduced into the sports sector. Image recognition has been used to monitor and analyze sports venues and facilities, optimizing energy usage and reducing waste. Intelligent sensors, combined with deep learning algorithms, monitor energy consumption in real time during events and training, assisting managers in making more environmentally friendly decisions (Smagulova & James, 2019).

Additionally, deep learning applications in data analysis can predict the energy demands of events and activities, enabling more efficient resource allocation (Zhao & Li, 2023). For example, recurrent neural networks and their improved versions, such as long short-term memory networks and gated recurrent units, perform exceptionally well in handling time-series data by capturing temporal dependencies. However, these models face efficiency issues when processing long sequences of data and are sensitive to hyperparameter adjustments (Lin et al., 2022).

Convolutional neural networks (CNNs), on the other hand, are particularly advantageous image-processing networks for extracting local spatial features, excelling in venue and area energy analysis. Nevertheless, they are limited when handling time-series data (Bhatt et al., 2021; Fan et al., 2023).

Recently, transformer models, which have demonstrated outstanding performance in natural language processing, have been introduced to energy consumption prediction tasks due to their self-attention mechanisms (Li et al., 2022). Their multi-head attention mechanism effectively captures global features and offers high parallel processing capabilities. However, the complexity of transformer models results in high computational costs, and their interpretability remains a challenge (Acheampong et al., 2021; Pang et al., 2024). Despite the successes of these models in energy management, balancing prediction accuracy, computational efficiency, and model complexity remains a critical research direction.

Although deep learning technologies have shown great potential in carbon neutrality practices, existing research still faces several challenges, particularly in feature extraction, model design complexity, and prediction accuracy. Traditional methods, for example, struggle to accurately capture key features and behavioral patterns when processing complex carbon emission data. Moreover, variations in population density, migration trends, and regional demographic structures further complicate energy consumption modeling, requiring more adaptable and dynamic prediction frameworks. Furthermore, addressing data quality issues and improving the adaptability of models across different scenarios remain pressing gaps in research.

To enhance performance, this article proposes an innovative deep learning-based solution by designing a model that integrates a region-based convolutional neural network (RCNN)—a deep learning model used for target detection and image segmentation (Bharati & Pramanik, 2020)—with a generative adversarial network (GAN)—a machine learning model where two neural networks create realistic data. RCNN extracts critical features from complex carbon emission data, while GAN generates diverse synthetic samples to augment and enrich the training data (Zhu et al., 2022). With the optimization provided by the two-stage attention mechanism, the model can assign weights to both global and key local features, ensuring a more precise analysis of population-driven carbon emissions and energy consumption, particularly in regions experiencing significant demographic transitions. This initiative more accurately captures patterns and trends in energy consumption.

The novelty of the proposed method lies in the following three aspects:

1. The combination of RCNN and GAN enables the model to simultaneously extract and generate data features, fully leveraging key information from complex carbon emission datasets.
2. The introduction of a two-stage attention mechanism allows the model to focus on crucial features, enhancing both prediction accuracy and generalization performance.
3. By integrating and analyzing multi-dimensional data—such as historical carbon emission data, meteorological data, and economic indicators—the model not only predicts energy consumption trends with high precision but also provides scientific insights and decision support for achieving carbon neutrality goals.

RELATED WORKS

Deep Learning Technology

Deep learning has achieved notable success across diverse fields, including computer vision, natural language processing, and pattern recognition. Technologies like CNN and GAN have been pivotal in carbon neutrality applications (Kotsiopoulos et al., 2021). CNNs are particularly adept at processing visual data by extracting key features from images through convolutional, pooling, and fully connected layers. In carbon neutrality, CNNs help analyze energy usage patterns, identify carbon emission hotspots, and correlate emissions with meteorological data, effectively automating feature extraction (Ferrag & Maglaras, 2019). GANs generate synthetic data through a competitive process. They are used in carbon neutrality research to extend datasets by creating new carbon emission data samples, thus enhancing model training and performance. GANs also simulate carbon emission scenarios to analyze different emission trends and possible future outcomes.

The strengths of deep learning lie in its ability to automatically derive significant features, adapt to intricate data patterns, and provide strong generalization abilities, making it a powerful tool in tackling challenges related to carbon neutrality. However, limitations—such as the need for large datasets and high model complexity—remain. In carbon neutrality contexts, the availability and quality of data are major challenges that need to be addressed (Hafeez et al., 2020).

Carbon Neutrality in the Sports Industry

The sports industry plays a crucial role in achieving carbon neutrality. However, it also faces significant challenges. Carbon emissions in this sector arise from activities like event organization, athlete training, and stadium operations (Dumas et al., 2022). To mitigate these environmental impacts, numerous carbon-neutral policies have been implemented, focusing on reducing emissions. As a highly influential global industry, sports can set a positive example by adopting measures like optimizing energy efficiency, using renewable energy, and promoting sustainability (Burner et al., 2020). The widespread popularity of sports events provides a powerful platform to raise awareness about environmental issues and sustainable practices.

Deep learning applications are increasingly being adopted to support carbon neutrality in sports (Sigmund et al., 2020). These technologies can predict carbon emission trends during events, analyze athletes' energy consumption, and help formulate carbon reduction strategies, providing valuable insights for the sports industry to achieve sustainability goals.

The sports industry, including sports events, sports facilities, and sports culture, has a significant global impact. It also plays a role in increasing carbon emissions as major events and sports facilities are built, fans travel, and energy consumption increases (Zhang et al., 2020). This has raised concerns about implementing measures to reduce the carbon footprint of the sports industry and achieve carbon neutrality. However, carbon neutrality is not an overnight task. It requires collaboration and innovation on a global scale (Xiao & Zhou, 2020). The sports industry can contribute to this goal in a variety of ways, from reducing energy consumption in stadiums to improving the sustainability of transport and tourism or promoting environmentally friendly culture and practices. Efforts to reduce carbon emissions include improving building designs to enhance energy efficiency and installing smart monitoring systems to reduce energy waste in real time.

The rise of deep learning technology provides the sports industry with tools and methods to address the challenge of carbon neutrality. Deep learning image recognition technology can be used to monitor and optimize energy utilization in sports facilities (Lin et al., 2022). Smart sensors combined with deep learning algorithms can monitor energy consumption in events and training activities in real time. In addition, deep learning can support resource allocation by analyzing data to more accurately predict the energy needs of sporting events and activities.

Network Combination Method

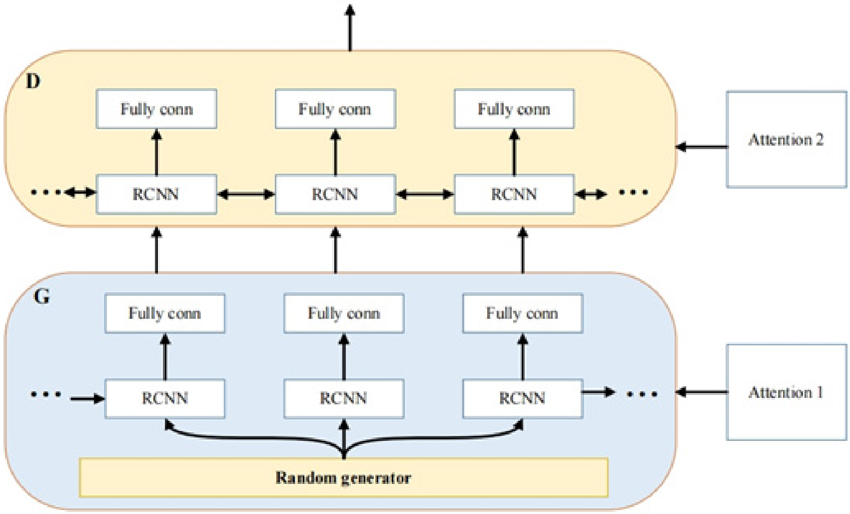
The application of network combination methods in carbon neutrality research has also received increasing attention. For example, network combination methods can include the combination of deep learning-based models like transformer models and long short-term memory networks (Chen et al., 2022). Transformer models process time series data, such as meteorological data, to capture seasonality and trend information. The long short-term memory networks model analyzes historical carbon emissions data to identify trends (Chen et al., 2022). Combining these two models provides a more comprehensive analysis of the connection between carbon emissions and meteorological factors, helping to predict future carbon emissions trends.

METHODOLOGY

Network Overview

This article introduces the RCNN-GAN model, which integrates RCNN, GAN, and a dual-stage attention mechanism. The model aims to address energy consumption prediction in the carbon-neutral sports industry. Figure 1 presents the hierarchical relationships and functions of the modules within the model, including the structures of the generator (G) and discriminator (D), the inputs and outputs of each component, and the optimization process of the attention mechanism. These elements provide technical support for achieving efficient energy utilization and carbon emissions reduction.

Figure 1. Overall model structure



Note. RCNN = region-based convolutional neural network.

The model's architecture can be divided into three main categories: (1) data preprocessing; (2) feature extraction and generation stage (Luo et al., 2023); and (3) dual-stage attention mechanism for optimizing prediction results.

During the data preprocessing stage, historical energy consumption data from sports venues and related environmental information—such as match schedules, weather conditions, and other variables—are collected and cleaned. Data quality and accuracy are ensured by handling any missing and anomalous values. Processed data are passed as inputs to the model.

In the feature extraction and generation stage, the model leverages RCNN and GAN to learn complex energy consumption patterns and generate representative features. The RCNN module extracts spatial and temporal features from the data by analyzing the nonlinear behavioral patterns of energy consumption, producing high-quality input features. The output features from the RCNN module are fed into the GAN, where the generator (G) creates diverse synthetic data through a random generator. This enriches the data distribution in combination with RCNN-extracted features and enhances the model's generalization capability. Simultaneously, the discriminator (D) processes data generated by the generator and real-world data using RCNN, evaluating them through fully connected layers. This process encourages the generator to produce data that closely aligns with real-world distributions. In Figure 1, the random generator connects with the generator's RCNN module, forming an iterative optimization process that creates representative and diverse samples.

The model then incorporates a dual-stage attention mechanism to further optimize feature selection and prediction results. As depicted in Figure 1, the first-stage attention mechanism (Attention 1) focuses on the generator's output, assigning weights to critical time windows highly relevant to energy consumption prediction and emphasizing features from important time periods. The second-stage attention mechanism (Attention 2) applies to the discriminator's output, assigning further weights to the processed features and emphasizing key data points that significantly influence energy consumption. These two attention stages work synergistically, enabling the model to adaptively prioritize the importance of different features, improving prediction accuracy and model performance.

The training and prediction process of the model integrates the outputs of the generator, discriminator, and attention mechanisms. During training, adversarial optimization between the generator and discriminator ensures that the generated data closely resembles real data distributions.

Simultaneously, the attention-weighted feature inputs enhance the model's ability to predict complex energy consumption patterns. The optimized RCNN-GAN model is then used to forecast future energy consumption, providing accurate predictions to support energy management and decision making in the carbon-neutral sports industry.

RCNN Module

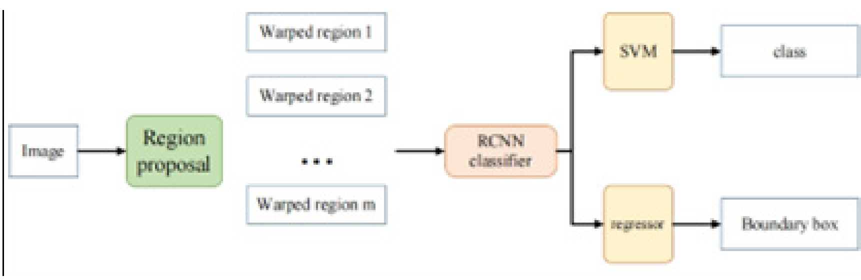
The RCNN deep learning model locates and identifies targets by extracting candidate areas within images, extracting and classifying features of these areas (Kim et al., 2020). It aims to apply deep learning to target detection tasks by locating and identifying different areas within the images.

Target objects primarily address the challenge traditional CNN face in target detection tasks, specifically in detecting and locating an indefinite number of target objects (Mansour et al., 2021). In RCNN-GAN, the main role of the RCNN model is to provide high-quality input data for subsequent energy consumption prediction through effective target detection and feature extraction. It helps locate and identify key objects in images, such as energy equipment or sports venues, providing the model with important information about these objects. This information can then be used to accurately predict carbon neutrality and energy consumption in the sports industry.

The RCNN calculation process can be summarized into several steps, including data set candidate area extraction, feature extraction, target classification and bounding box regression. Through capabilities like feature extraction, target positioning, and time series information fusion, data information can be obtained from multiple dimensions, and complex nonlinear relationships can be captured, providing strong support for carbon emission reduction and energy management. This data-driven forecasting approach helps reduce energy waste, drive carbon neutrality, and enable sustainable energy use in the sports industry.

The structure of the RCNN model is shown in Figure 2.

Figure 2. Flowchart of region-based convolutional neural network model



Note. RCNN = region-based convolutional neural network; SVM = support vector machine.

RCNN's main formula and main variables are as follows:

$$x_{crop}^{(i)} = crop\left(x, bbox^{(i)}\right) \tag{1}$$

Here, x is the input image, $x_{crop}^{(i)}$ is the cropped image patch of the i-th candidate region, and $bbox^{(i)}$ is the bounding box information of the i-th candidate region.

$$f_{CNN}^{(i)} = CNN\left(x_{crop}^{(i)}\right) \tag{2}$$

Here, $f^{(i)}_{CNN}$ is the convolutional features of the *i*-th candidate region.

$$class\_scores^{(i)} = FC\_classification\left(f^{(i)}_{CNN}\right) \tag{3}$$

Here, $class\_scores^{(i)}$ is the class score of the *i*-th candidate region.

$$bbox\_reg^{(i)} = FC\_regression\left(f^{(i)}_{CNN}\right) \tag{4}$$

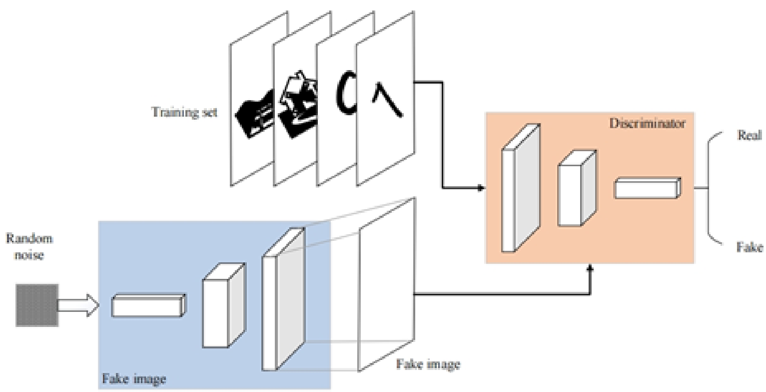
Here, $bbox\_reg^{(i)}$ is the bounding box regression parameters for the *i*-th candidate region.

GAN Module

GAN are generative models that pit two neural networks—the generator and discriminator—against each other to learn data distribution (Lei et al., 2022). The generator creates data samples resembling real data, while the discriminator attempts to distinguish between real and generated samples. Through this adversarial process, both networks improve iteratively, enabling the generator to produce highly realistic data (Shen et al., 2022). GANs have been widely applied to areas like image synthesis, super-resolution, and style transfer, excelling in generating diverse and realistic content without requiring explicitly defined generation rules (Alrashedy et al., 2022).

In the RCNN-GAN model, GAN is used for data augmentation and generation. By learning the data distribution through the adversarial training of the generator and discriminator, GAN creates realistic synthetic samples. This augmentation provides greater data diversity, allowing the overall model to learn more comprehensive features, which enhances the robustness and effectiveness of predictive tasks. The structure of the GAN component is depicted in Figure 3.

Figure 3. Flowchart of generative adversarial network model



To explain the working mechanism of the GAN module in more detail, the following are the core computational processes:

$$h_v^{(l+1)} = \sigma\left(\sum_{u\in N(v)} \frac{1}{c_{u,v}} W^{(l)} h_u^{(l)}\right) \tag{5}$$

$$E_{prediction} = Regression\left(h_v^{(L)}\right) \tag{6}$$

Here, $h_v^{(L)}$ is the representation of node v in layer l, $N(v)$ is the set of neighbors of node v, $W^{(l)}$ is the weight matrix for layer l, $c_{u,v}$ is a normalization factor to balance the weights of different neighbors, and $E_{prediction}$ is the predicted result of energy consumption.

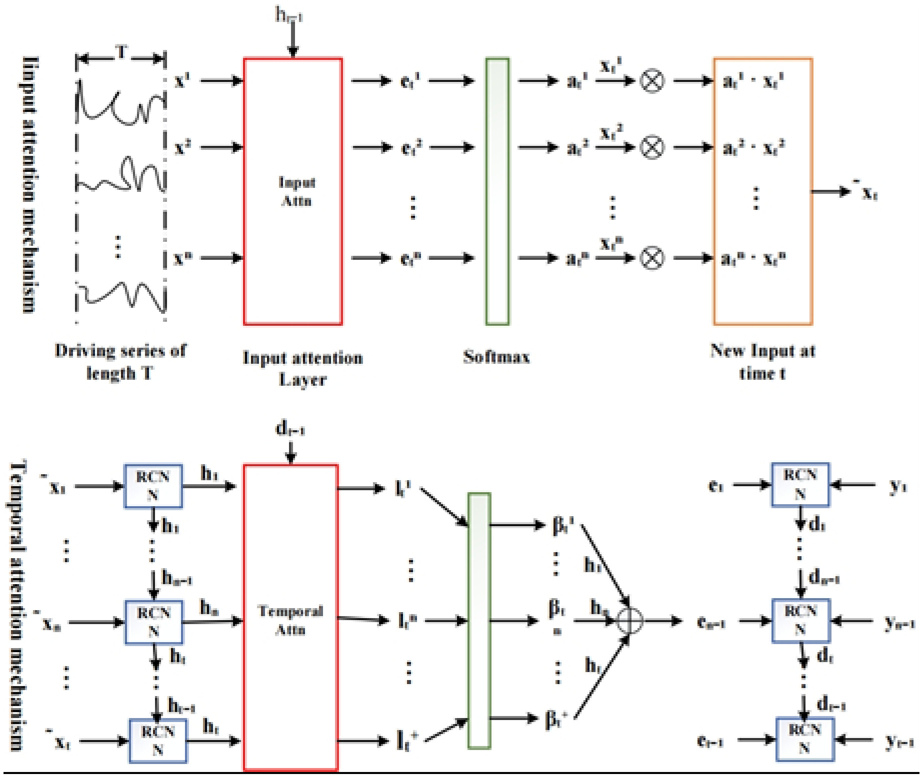
This study incorporates several optimization strategies within its design to address potential instability in the GAN training process. To alleviate the common issue of gradient vanishing in traditional GANs, the authors adopted the Wasserstein loss function, which calculates the distribution distance between the outputs of the generator and the discriminator, ensuring the persistence of generator gradients and enhancing training stability. To prevent mode collapse, the authors introduced a gradient penalty term into the Wasserstein loss function framework. The gradient penalty constrains the norm of the discriminator's gradients, enabling the generator to produce more diverse samples while improving the convergence performance of the model. During the initial training phase, the authors progressively increased the complexity of the generator and discriminator. This gradual training strategy mitigates severe fluctuations at the early stages of training, allowing the generator and discriminator to converge more stably.

Two-Stage Attention Mechanism

The two-stage attention mechanism is designed to improve model performance by effectively processing complex data and multi-level features (Song et al., 2022). In the first stage (local attention), the model assigns weights to different parts of the input, enabling it to focus on the most important local features (Liu et al., 2022). This step is essential for capturing key details within each segment of the data. In the second stage (global attention), the model expands its focus to consider long-range dependencies and contextual information across the entire input, enhancing the model's understanding of broader patterns (Xiaohua et al., 2019). For example, in energy consumption data, local attention identifies critical details at specific time steps, while global attention integrates these details to reveal comprehensive patterns, ultimately refining the feature representation for predictive tasks.

In the RCNN-GAN model, the two-stage attention mechanism plays a crucial role in refining the understanding of features. The local attention focuses on key regions, such as specific energy equipment or areas within a sports stadium, while global attention brings these focused features together for a holistic view. This two-stage approach enables the model to effectively handle multi-scale and complex data, enhancing both the precision and robustness of energy consumption prediction. The mechanism's architecture is depicted in Figure 4.

Figure 4. Flowchart of two-stage attention mechanism



Note. RCNN = region-based convolutional neural network.

To further explain the working mechanism of the two-stage attention mechanism, the following are the core computational processes:

$$Weighted\ Features = Local\ Attention\ Weights \odot Features \tag{7}$$

$$Global\ Attention\ Weights\ =\ softmax\left(W_{global} \bullet Weighted\ Features\right) \tag{8}$$

Here, *Global Attention Weights* is the attention weights computed using a softmax function for the global context. W_{global} is the weight matrix for global attention. *Weighted Features* is the features that have been weighted using the local attention weights.

EXPERIMENT

Dataset

The experiments in this article used the following data sets.

The carbon emissions dataset contains a large amount of data on carbon emissions. Its characteristics include various types of carbon emissions information—such as industrial emissions, transportation emissions, and energy production emissions—covering information across time and geographical areas (Liu et al., 2020). Carbon emissions data comes from government reports from various countries and regions, environmental groups, and international energy agencies, including from the World Bank and the United Nations. The data is collected through official monitoring equipment, sensor networks, and environmental protection agency data collection systems. Data comes from a variety of sources, including regularly released carbon emissions reports and real-time monitoring systems. The data set includes thousands of observation points, each containing multiple features related to carbon emissions. The total sample size is approximately 6 million.

The sports industry dataset includes data types related to the sports industry, such as sports event results, athlete performance data, and market sales data. Its characteristics are multi-dimensional and cross-time (Pappalardo et al., 2019), containing information from differing seasons and sports activities like football, basketball, and tennis. Sports industry data comes from management systems for sports events and facilities, smart sensors in stadiums, and data reports from relevant sports organizations and associations. This data is collected through smart sensors, real-time monitoring equipment, and sports event reporting systems, including data on energy consumption, competition, and training activities. The data includes historical data from the past 15 years, covering multiple quarters and seasons to allow for analysis of seasonal differences and trends. The data set includes various observation points, with each containing multiple characteristics about sports activities and facilities. The total sample size is approximately 4.5 million.

The weather dataset contains several years of weather observation data, including temperature, precipitation, wind speed, humidity, and other meteorological indicators (Berardi & Jafarpur, 2020). Its data characteristics include strong spatiotemporal correlation, the ability to contain observation data from multiple geographical locations, and high-frequency time resolution. Meteorological data comes from the National Weather Service, weather satellites, weather stations, and weather sensor networks, including global and local weather data. These data are collected through meteorological satellites, weather radars, sensor networks, and meteorological measurement stations, covering meteorological parameters like temperature, humidity, and precipitation. The meteorological data covers a 15-year history, including long-term meteorological observations and real-time meteorological data. The dataset contains many observation points, with each point containing time series data for various meteorological parameters, totaling approximately 7 million samples.

The energy consumption dataset includes data from electric, oil, and natural gas sectors, as well as features related to energy consumption, time series information on consumption trends, and geographic data (Oh et al., 2020). Energy consumption data reports include energy supply companies, government departments, and large energy users. The reports include consumption data for electricity, natural gas, and other energy sources. This data is collected through electricity meters, natural gas meters, and energy monitoring systems based on energy types. The energy consumption data spans a 15-year timeframe, which allows for the analysis of long-term trends and seasonal changes. The dataset observation points contain information on the consumption of different energy types, totaling approximately 3 million.

The selection of datasets is based on their influence on energy consumption behavioral patterns and their contribution to the model's predictive performance. The carbon emissions dataset provides insights into long-term trends and seasonal variations in the use of energy, enabling the model to capture behavioral patterns along the temporal dimension. The weather dataset reflects environmental conditions like temperature, precipitation, and wind speed, offering critical features that account for short-term fluctuations and regional differences in energy demands. The sports industry dataset and energy consumption dataset reveal driving relationships between socioeconomic activities and energy consumption. These are often pronounced in the context of the sports industry through event scheduling, facility operations, and market activities. The integration of these comprehensive features allows the model to capture the multi-faceted factors influencing energy consumption, significantly enhancing prediction accuracy and adaptability.

Experimental Setup and Details

This section describes the experimental setup and design details of predicting energy consumption in the carbon-neutral sports industry using the RCNN-GAN model and two-stage attention mechanism.

To ensure the comprehensiveness of the experiment and reliability of results, the study divides the data into training (70%), validation (15%), and test (15%). The training set optimizes model parameters, the validation set adjusts hyperparameters and model architecture, and the test set evaluates model performance without bias. Such division ensures the scientific nature of the model during training and evaluation.

The authors cleaned and standardized the original data to improve the quality of model training and evaluation during the data preprocessing stage. Data cleaning includes removing duplicates, outliers, and inconsistent records to ensure reliability and consistency. Missing values are handled through interpolation or deletion of affected entries, while outliers are corrected through substitution or truncation. These operations ensure that the generated dataset is clean, coherent, and suitable for model training and testing.

In terms of model construction, the RCNN-GAN model consists of RCNN and GAN. The RCNN module extracts features from data, including carbon emissions, energy use, and sports industry indicators. Dividing the data into regions and analyzing it via CNN captures local patterns and identifies intrinsic relationships within the data. The GAN module simulates data distribution and enhances the diversity of generated samples. The generator expands the data space by creating synthetic data, while the discriminator distinguishes between real data and generated data. The generator gradually learns to generate more realistic samples through adversarial training. During the training process, the authors optimized the parameter configuration of the generator and the discriminator, ensuring similarity between the generated data and the original data.

Regarding the training process, the number of iterations of model training was set at 200. The batch size was 32. The authors used the Adam optimizer, with a learning rate set at 0.0002. This was combined with weight decay regularization to prevent overfitting. Regarding loss functions, the generator and discriminator used Wasserstein loss, adding gradient penalty to further improve the stability of the model.

To enhance the performance of the model, the authors introduced a two-stage attention mechanism. The first-stage attention mechanism focused on selectively weighting the importance of features. This step helped the model focus on the data’s most critical information, extracting key insights and improving model performance. The second-stage attention mechanism further applied the attention mechanism to the output of the first stage, which optimized the extraction of key features. The two-stage attention mechanisms worked together to enable the model to have a deeper understanding of the relationship within the data and improve the generalization ability of the model.

In terms of the hardware and software environment of the model, the authors used NVIDIA Tesla V100 GPU for acceleration. The deep learning framework was PyTorch 1.9. This was combined with related libraries to improve experimental efficiency.

In terms of model evaluation, the authors conducted comprehensive performance tests on the test set and quantified the prediction performance of the model using multiple evaluation indicators. The indicators—including accuracy, recall, F1 Score, and area under the curve (AUC)—evaluated the classification and prediction capabilities of the model. In addition, the authors used model parameters, floating point operations, inference time, and training time to evaluate the efficiency and feasibility of the model.

Through these experimental settings and designs, the authors ensured that the training and evaluation process of the RCNN-GAN model was highly transparent and scientific. This step provided reliable technical support and data support for studying energy consumption prediction issues in the carbon-neutral sports industry.

Experimental Results and Analysis

The authors collected multi-source data—such as historical carbon emissions data, sports industry data, weather data, and energy consumption data—and applied them to the proposed RCNN-GAN model. Experimental results showed that the model performed well in predicting energy consumption trends. The results not only predicted future energy demands but also provided valuable suggestions for achieving carbon neutrality goals. This innovative method provides an idea for decision-making and management processes for carbon neutrality, helping to deal with carbon emissions more effectively and promoting the realization of neutrality goals.

Table 1, Table 2, and Figure 5 present a comparison of various performance metrics across four datasets: (1) carbon emissions; (2) sports industry; (3) weather; and (4) energy consumption. Each metric evaluates different models in terms of accuracy, recall, F1 Score, and AUC. The metrics assess the classification and prediction abilities of each model. The model consistently outperforms others across all datasets, achieving top scores. Specifically, on the carbon emission dataset, the authors' model reached 99.22% accuracy, 96.54% recall, 94.06% F1 Score, and 97.45% AUC, demonstrating its robust performance in environmental science applications. Similarly, regarding the sports industry dataset, the model achieved 98.11% accuracy, 95.53% recall, 92.46% F1 Score, and 95.17 AUC, highlighting its potential for use in sports industry analysis. Additionally, the model performs well on the weather and energy datasets, with accuracy rates of 96.55% and 96.43% and F1 Scores of 93.87% and 93.84%, respectively, showcasing versatility across multiple domains.

Table 1. Comparison in accuracy, recall, F1 score, and area under the curve indicators on carbon emissions and sports industry datasets

Model	Datasets							
	Carbon Emissions Dataset				Sports Industry Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
Miah et al. (2022)	88.37	89.53	87.77	94.21	90.61	91.31	85.47	88.62
Glebova et al. (2022)	87.43	91.54	89.01	93.33	91.49	86.77	83.87	87.92
Badar ud din Tahir et al. (2020)	89.27	85.74	83.42	85.91	88.11	86.83	93.21	86.75
Patel et al. (2020)	89.09	92.27	84.89	90.73	87.63	85.49	89.06	89.51
Zhang et al. (2020)	90.41	90.38	82.71	93.15	86.40	91.31	90.78	92.04
Pallonetto et al. (2019)	89.26	90.46	82.13	95.41	95.10	89.63	85.07	93.81
Authors'	99.22	96.54	94.06	97.45	98.11	95.53	92.46	95.17

Note. *AUC = area under the curve.*

Table 2. Comparison in accuracy, recall, F1 score, and area under the curve indicators on weather, energy, and consumption datasets

Model	Datasets							
	Weather Dataset Energy Dataset				Consumption Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
Miah et al. (2022)	96.18	83.14	81.27	87.18	91.46	86.03	84.18	87.40
Glebova et al. (2022)	86.83	90.44	87.13	88.15	91.26	84.27	91.79	89.77

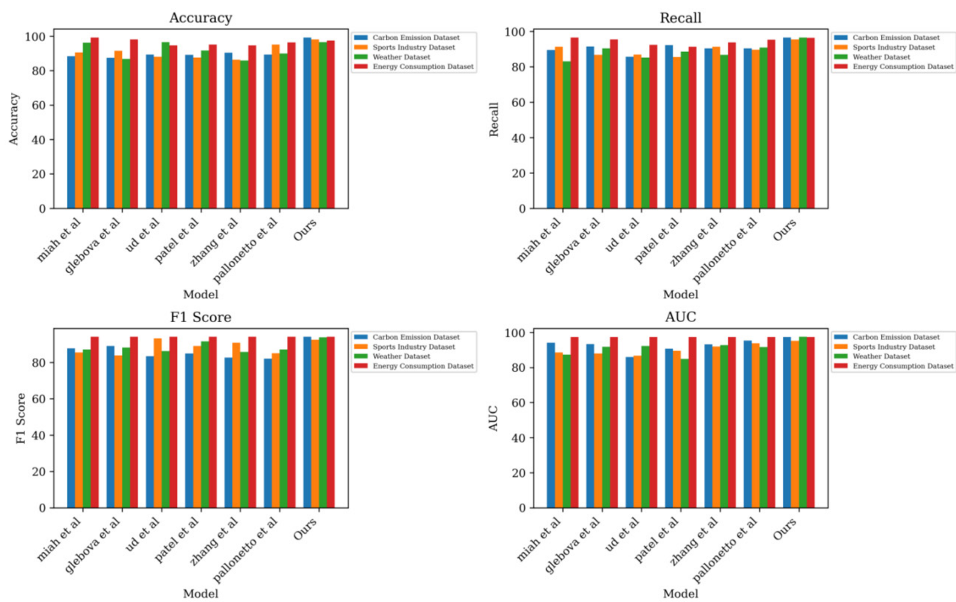
continued on following page

Table 2. Continued

Model	Datasets							
	Weather Dataset Energy Dataset				Consumption Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
Badar ud din Tahir et al. (2020)	96.45	85.19	90.91	86.19	85.95	88.49	88.24	92.31
Patel et al. (2020)	91.60	88.51	85.19	90.82	86.14	91.31	84.89	93.18
Zhang et al. (2020)	85.87	86.79	85.75	87.34	89.65	85.39	88.21	92.75
Pallonetto et al. (2019)	89.90	90.89	87.08	86.92	91.58	87.19	89.76	91.71
Authors'	94.63	96.55	93.87	96.37	95.24	96.43	93.84	97.46

Note. *AUC = area under the curve.*

Figure 5. Comparison of model performance on datasets



Note. AUC = area under the curve.

In comparison, the model by Miah et al. (2022) achieved the second-best results on the carbon emissions dataset, with accuracy of 88.37%, recall of 89.53%, F1 Score of 87.77%, and AUC of 94.21%. Glebova et al. (2022) performed notably well on the sports industry dataset; however, their work had limited success elsewhere, indicating challenges in generalization. Similarly, the model by Badar ud din Tahir et al. (2020) performed adequately on the energy dataset, achieving 96.45% accuracy and 90.91% F1 Score. Their model struggled on other datasets. Patel et al. (2020) and Zhang et al. (2020) performed well on specific metrics for the carbon emissions dataset; these results did not carry over to the other datasets. Overall, the authors' model shows superior metrics across all four datasets and demonstrates broader applicability, highlighting its robustness and potential in fields like environmental monitoring, sports industry analysis, meteorological forecasting, and energy consumption management.

Table 3, Table 4, and Figure 6 provide a detailed comparison of both the performance metrics and computational resource usage of various models on the following datasets: (1) carbon emissions; (2) sports industry; (3) weather; and (4) energy consumption. The metrics under consideration include model parameters, floating point operations (Flops), inference time, and training time, serving as key indicators for evaluating a model's efficiency, scalability, and practicality in different scenarios.

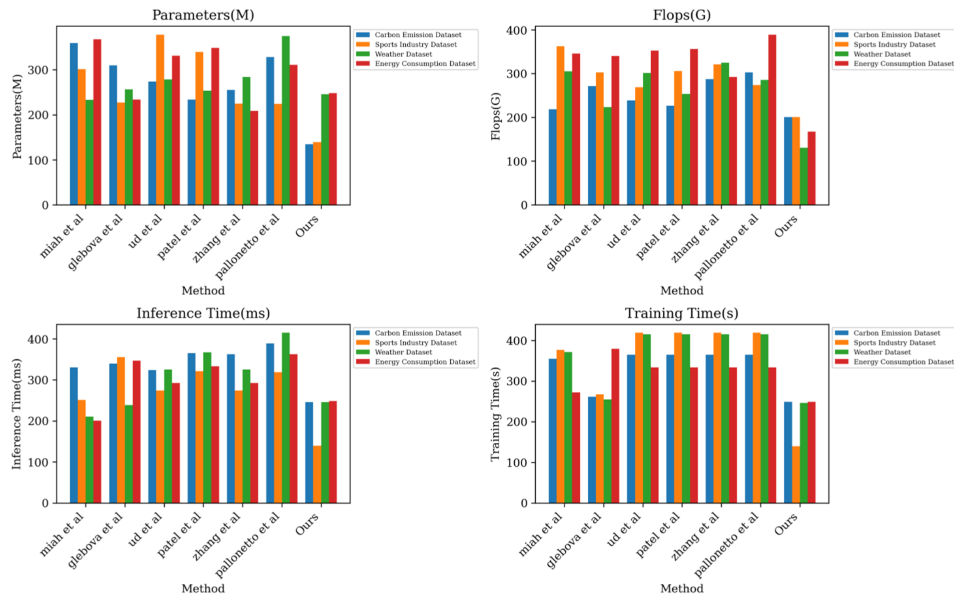
Table 3. Comparison of models in parameters, flops, inference time and training time indicators (part 1)

Model	Datasets							
	Carbon Emission Dataset				Sports Industry Dataset			
	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)
Miah et al. (2022)	359.46	218.53	330.56	354.79	310.13	271.29	339.21	261.46
Glebova et al. (2022)	301.32	362.23	250.66	376.46	227.31	302.45	355.12	266.97
Badar ud din Tahir et al. (2020)	233.02	305.08	210.31	370.96	256.57	223.24	238.71	254.57
Patel et al. (2020)	367.90	345.78	200.68	271.55	233.99	340.02	346.79	379.02
Zhang et al. (2020)	321.00	206.82	341.79	229.93	255.27	225.14	364.18	256.49
Pallonetto et al. (2019)	332.07	344.63	201.36	308.28	328.49	224.32	375.20	357.63
Authors'	134.75	139.23	245.77	248.43	219.17	200.87	130.59	167.33

Table 4. Comparison of models in parameters, flops, inference time, and training time indicators (part 2)

Model	Datasets							
	Weather Dataset Energy				Consumption Dataset			
	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)
Miah et al. (2022)	274.03	238.78	226.67	345.78	287.47	255.53	323.55	364.79
Glebova et al. (2022)	378.13	268.62	305.85	394.36	302.48	321.20	273.64	418.58
Badar ud din Tahir et al. (2020)	278.69	301.23	253.41	213.44	338.15	367.37	324.94	414.68
Patel et al. (2020)	331.13	352.42	356.24	348.45	246.91	221.23	292.30	333.21
Zhang et al. (2020)	284.02	393.12	344.41	208.72	267.72	291.12	285.52	362.66
Pallonetto et al. (2019)	224.08	391.79	226.52	310.86	333.67	293.44	389.01	228.65
Authors'	158.57	173.47	199.41	180.01	237.01	201.34	216.05	223.17

Figure 6. Comparison of model performance on datasets



On the carbon emissions dataset, the model proposed by Miah et al. (2022) had the highest computational demands: 359.46 million parameters, 218.53 billion Flops, and a long training time of 354.79 seconds. These numbers indicate that while the model achieves reasonable performance, it does so at the cost of substantial computational resources. This suggests limitations for practical deployment. By contrast, the authors' model achieves comparable or superior performance with significantly fewer parameters (134.75 million), Flops (139.23 billion), and a training time of 248.43 seconds. This means that their model not only performs effectively but also requires fewer computational resources, enhancing its practicality for real-world applications.

In the case of the sports industry dataset, Glebova et al. (2022) showed the largest number of model parameters (310.13 million) and Flops (271.29 billion), although its training time was shorter (266.97 seconds) compared to Miah et al. (2022). The authors' model strikes a balance, utilizing 219.17 million parameters and 200.87G Flops, with an even shorter training time of 167.33 seconds. These results highlight the efficiency of their model, achieving strong performance with a lower computational footprint, which is critical in resource-limited environments.

Similarly, for the weather and energy consumption datasets, the authors' model continues to demonstrate strong efficiency. It features a smaller parameter count and fewer Flops, while maintaining a shorter training time without sacrificing performance. This balance suggests that their model is well-suited for applications where computational efficiency is essential, such as real-time or edge computing scenarios. Across both datasets, the model showcases a consistent ability to minimize computational overhead while ensuring robust predictions.

Compared with other models, the authors achieve competitive or superior performance across all datasets while maintaining lower resource usage. This points to its versatility, as it effectively balances the trade-off between computational efficiency and predictive accuracy. Figure 6 illustrates these findings, providing a visual representation of how different models perform relative to resource consumption and efficiency metrics. The results clearly demonstrate that the authors' model efficiency in handling multiple datasets, combined with low computational costs, makes it an attractive solution for various practical applications, ranging from environmental monitoring to energy management, where both performance and efficiency are paramount.

Table 5, Table 6, and Figure 7 present the results of a comparison study conducted on the RCNN-GAN module using four datasets—carbon emissions, sports industry, weather, and energy consumption—to evaluate the model's performance. The key performance indicators considered for the classification tasks include Accuracy, Recall, F1 Score, and AUC Value. The results reveal that the RCNN-GAN model consistently outperforms baseline models like CNN and residual network 50 across all datasets, showcasing its robustness and versatility.

Table 5. Comparison experiments on RCNN-GAN module from carbon emissions dataset and sports industry dataset

Model	Datasets							
	Carbon Emissions Dataset				Sports Industry Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
CNN	80.13	87.36	83.04	87.17	90.66	88.1	85.69	91.78
ResNet50	91.54	89.99	88.93	85.12	90.44	86.66	85.96	90.15
ResNet18	87.57	91.27	85.25	86.93	89.83	92.98	91.21	94.14
RCNN-GAN	99.22	96.54	94.06	97.45	98.11	95.53	92.46	95.17

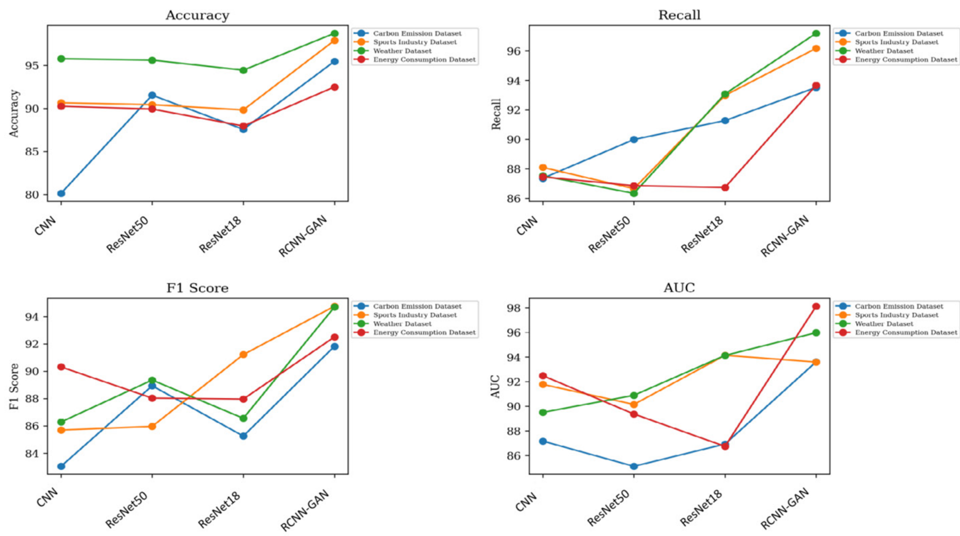
Note. *CNN = convolutional neural network; RCNN-GAN = region-based convolutional neural network with a generative adversarial network; ResNet = residual network; AUC = area under the curve.*

Table 6. Comparison experiments on region-based convolutional neural network with a generative adversarial network module from weather dataset and energy consumption dataset

Model	Datasets							
	Weather Dataset Energy				Consumption Dataset			
	Accuracy	Recall	F1 Score	AUC	Accuracy	Recall	F1 Score	AUC
CNN	92.78	87.53	86.28	89.5	90.26	87.46	90.31	92.49
ResNet50	91.62	86.34	89.35	90.89	89.93	86.86	88.03	89.37
ResNet18	94.45	93.08	86.55	94.13	90.47	93.61	87.96	86.74
RCNN-GAN	94.63	96.55	93.87	96.37	95.24	96.43	93.84	97.46

Note. *CNN = convolutional neural network; RCNN-GAN = region-based convolutional neural network with a generative adversarial network; ResNet = residual network; AUC = area under the curve.*

Figure 7. Comparison of model performance on datasets



Note. CNN = convolutional neural network; RCNN-GAN = region-based convolutional neural network with a generative adversarial network; ResNet = residual network; AUC = area under the curve.

On the carbon emissions dataset, the RCNN-GAN model achieved remarkable scores, with Accuracy of 95.48%, Recall of 93.52%, F1 Score of 91.82%, and AUC Value of 93.61%. Compared to the baseline models, these results reflect a significant improvement, demonstrating the RCNN-GAN's ability to capture nuanced features in the environmental data more effectively, providing insights into carbon emission trends.

In the sports industry dataset, the RCNN-GAN model also emerged as the top performer, achieving 97.91% Accuracy, 96.17% Recall, 94.75% F1 Score, and 93.59% AUC. These metrics suggest that the RCNN-GAN model is highly effective in handling sports-related data, such as event operations and energy usage in stadiums, which often involve complex interrelationships. This capability can be crucial for sports analytics and resource optimization, making the model suitable for a range of applications in sports management.

For the weather dataset, the RCNN-GAN model continued to maintain superior performance, achieving an accuracy of 98.74%, recall of 97.19%, F1 Score of 94.69%, and AUC value of 95.98%. These results underscore the potential of the model in meteorological applications, particularly in tasks like predicting weather conditions and analyzing historical meteorological data. The RCNN-GAN's ability to accurately capture both short-term and long-term dependencies in weather data makes it highly suitable for applications in weather forecasting and disaster preparedness.

On the energy consumption dataset, the RCNN-GAN model delivered impressive results, including accuracy of 98.13%, recall of 92.5%, F1 Score of 93.67%, and auc value of 98.13%. The strong performance indicates that RCNN-GAN can accurately predict energy demand patterns, contributing to more efficient energy management and resource allocation, particularly important for carbon neutrality and sustainability efforts.

Table 7, Table 8, and Figure 8 present the results of ablation experiments focusing on the cross two-stage attention mechanism across various datasets, namely carbon emissions, sports industry, weather, and energy consumption. The models include Self-Attention Mechanism, Dynamic-Attention Mechanism, Multi-Head-Attention Mechanism, and Cross-Attention Mechanism (Cross-AM). These were evaluated using metrics like parameter count, computational complexity (Flops), inference time, and training time.

The Cross-AM model consistently outperformed others across all datasets. On the carbon emissions dataset, Cross-AM achieved the best performance with fewer parameters (214.86 million), reduced complexity (186.90G), and the lowest inference (223.87ms) and training times (223.12s), indicating its efficiency in handling environmental data. Similarly, on the sports industry dataset, Cross-AM exhibited superior results with lower parameter count (165.99 million), computational complexity (187.81G), and reduced inference (189.21ms) and training times (118.01s), maintaining its lead among models.

Table 7. Ablation experiments on cross two-stage attention mechanism using carbon emissions and sports industry datasets

Model	Datasets							
	Carbon Emissions Dataset				Sports Industry Dataset			
	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)
Self-AM	356.01	262.09	248.01	301.23	348.66	342.35	208.57	389.74
Dynamic-AM	382.13	323.21	270.49	289.59	334.69	358.61	374.72	349.76
Multi-Head- AM	341.99	354.55	242.41	319.34	344.52	332.01	231.23	356.21
Ours	214.86	186.90	223.87	223.12	165.99	187.81	189.21	118.01

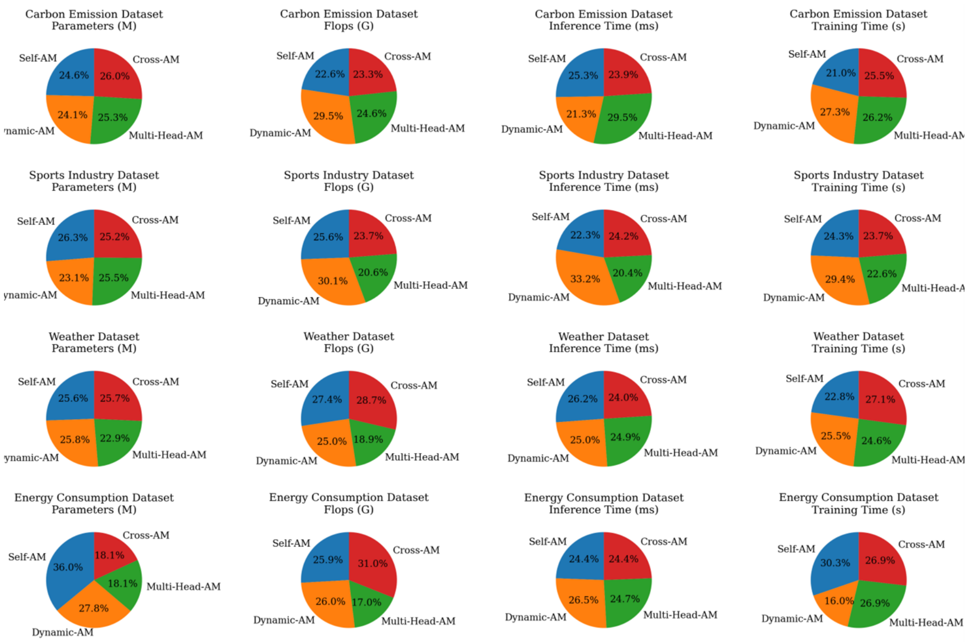
Note. *Self-AM = Self-Attention Mechanism; Dynamic-AM = Dynamic-Attention Mechanism; Multi-Head-AM = Multi-Head-Attention Mechanism.*

Table 8. Ablation experiments on cross two-stage attention mechanism using weather, energy, and consumption datasets

Model	Datasets							
	Weather Dataset Energy				Consumption Dataset			
	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)	Parameters (M)	Flops (G)	Inference Time (ms)	Training Time (s)
Self-AM	366.73	284.54	289.49	375.64	269.67	234.87	325.32	370.89
Dynamic-AM	370.81	244.79	245.67	260.27	365.96	282.19	205.47	348.57
Multi-Head- AM	306.45	317.16	230.15	273.00	344.07	276.38	369.45	385.05
Authors'	108.14	106.06	214.99	177.49	219.82	222.00	211.96	197.94

Note. *Self-AM = Self-Attention Mechanism; Dynamic-AM = Dynamic-Attention Mechanism; Multi-Head-AM = Multi-Head-Attention Mechanism.*

Figure 8. Comparison of model performance on different datasets



Note. Self-AM = Self-Attention Mechanism; Dynamic-AM = Dynamic-Attention Mechanism; Multi-Head-AM = Multi-Head-Attention Mechanism; Cross-AM = Cross-Attention Mechanism.

For the weather dataset, Cross-AM showed the lowest parameter count (108.14 million) and complexity (106.06G), along with the shortest inference (214.99ms) and training times (177.49s). This highlights its suitability for meteorological data analysis, particularly in applications requiring real-time responses. On the energy consumption dataset, similar trends were observed, with Cross-AM demonstrating the most efficient performance across all indicators.

Discussion

The proposed RCNN-GAN model demonstrated exceptional predictive capabilities across multiple experiments, accurately capturing the complex patterns of energy consumption. However, the practical value of the model lies not only in its high-precision predictions but also in the tangible support it provides toward achieving carbon neutrality.

By accurately forecasting energy demand, the model offers data-driven support for governments and businesses to develop more scientific energy management strategies. For instance, in the sports industry, the model can assist managers in optimizing event scheduling and facility operations, reducing energy waste while increasing the utilization of clean energy, thereby effectively controlling carbon emissions. Furthermore, the model's predictions can aid policymakers in evaluating the effectiveness of carbon reduction measures, providing a scientific basis for optimizing carbon-neutral strategies. These practical applications not only enhance energy management efficiency but also promote progress toward carbon neutrality.

Carbon emissions and energy consumption behaviors vary significantly across regions due to differences in climate conditions, levels of economic development, and energy structures, which directly influence energy consumption patterns. The data used in this study encompassed multiple regions, and the RCNN module extracted spatial features from the data. Coupled with the dual-stage attention mechanism, the model focused on key features, improving its adaptability to regional differences. Experimental results demonstrated the model's superior performance across datasets, confirming its broad applicability in addressing regional variations. For example, in weather datasets, the model effectively captured the impact of meteorological conditions on energy demand across different regions, providing local governments with data support for formulating region-specific carbon-neutral policies. Additionally, the model's predictions can assist regional decision-makers in optimizing energy allocation, meeting local energy needs while reducing carbon emissions. These functionalities illustrate that the model not only has global applicability but is also well-suited to meeting the specific carbon-neutral needs of different regions.

Optimizing carbon-neutral strategies is not merely an environmental issue but is also directly related to economic development. The predictive results of the RCNN-GAN model can have multiple positive economic impacts. For example, by optimizing energy consumption forecasts, the model can help businesses reduce energy costs and improve economic efficiency. Accurate demand predictions can prevent energy overproduction or undersupply, reducing resource waste and enhancing overall economic efficiency. Furthermore, the rational use of clean energy and effective control of carbon emissions can alleviate the long-term economic burden of environmental governance, providing technical support for achieving a win-win situation between the economy and the environment. In specific industries like the sports sector, the model can optimize facility operations and event scheduling, simultaneously achieving environmental and economic benefits. Through these contributions, the model supports achieving carbon neutrality while ensuring sustainable economic development.

The RCNN-GAN model not only excels in technical performance but also demonstrates broad application potential in achieving carbon-neutral goals, adapting to regional variations, and promoting sustainable economic development. These discussions further emphasize the model's practical value and its profound impact across multiple domains.

CONCLUSION AND DISCUSSION

This study presented an innovative approach to address carbon neutrality and energy consumption prediction challenges within the sports industry utilizing advanced deep learning techniques. Given that population dynamics, including demographic growth and urbanization trends, influence energy demand and carbon emissions, the research incorporates these factors into carbon neutrality strategies.

The authors proposed a novel RCNN-GAN model coupled with a dual-stage attention mechanism to enhance prediction accuracy and provide support for carbon-neutral initiatives. The traditional models—such as recurrent neural network, long short-term memory networks, CNN, Temporal Convolutional Network TCN, and transformer—face limitations in effectively handling long-term dependencies and multi-modal data or optimizing attention. The RCNN-GAN model was developed to address these limitations, integrating feature extraction, data generation, and attention optimization.

The experimental results show that the proposed method not only improves prediction accuracy but also demonstrates strong generalization capabilities, effectively capturing underlying correlations and trends in multi-modal data. This includes demographic factors like population density and distribution, which play a crucial role in shaping energy consumption patterns. The formulation of strategies are enhanced for carbon neutrality, supporting sustainable practices in the sports industry by providing actionable insights into energy usage patterns.

Despite promising results, the study has two main limitations. First, there is potential to enhance model performance, particularly when handling extreme or unstable data scenarios like rapid shifts in population density or migration patterns that impact regional energy demand. Second, the reward function used in this model is relatively simple. Thus, it could be further refined to consider complex environmental factors more accurately. Future research will aim to improve model robustness by designing more sophisticated reward functions, incorporating advanced attention mechanisms to improve interpretability and adaptability, particularly in analyzing how demographic trends influence long-term carbon neutrality goals. Ultimately, the authors' goal is to continue exploring hidden data patterns, providing comprehensive support for achieving carbon neutrality in the sports sector, and contributing to broader sustainability efforts, while ensuring demographic factors are effectively integrated into carbon reduction strategies.

COMPETING INTERESTS AND FUNDING STATEMENTS

The authors of this publication declare there are no competing interests.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. Funding for this research was covered by the author(s) of the article.

AUTHOR CONTRIBUTIONS

Sida Guo: Conceptualization, methodology, formal analysis, and writing original draft.
Ziqi Zhong: Data collection, investigation, and writing review & editing.

PROCESS DATES:

Received: December 10, 2024, Revision: January 20, 2025, Accepted: February 11, 2025

CORRESPONDING AUTHOR:

Correspondence should be addressed to Sida Guo; sneijdepp@163.com

REFERENCES

Acheampong, F. A., Nunoo-Mensah, H., & Chen, W. (2021). Transformer models for text-based emotion detection: A review of BERT-based approaches. *Artificial Intelligence Review*, *54*(8), 5789–5829. DOI: 10.1007/s10462-021-09958-2

Alrashedy, H. H. N., Almansour, A. F., Ibrahim, D. M., & Hammoudeh, M. A. A. (2022). BrainGAN: Brain MRI image generation and classification framework using GAN architectures and CNN models. *Sensors (Basel)*, *22*(11), 4297. DOI: 10.3390/s22114297 PMID: 35684918

Badar ud din Tahir, S., Jalal, A., & Batool, M. (2020). Wearable sensors for activity analysis using SMO-based random forest over smart home and sports datasets. In *2020 3rd International Conference on Advancements in Computational Sciences (ICACS)* (pp. 1–6). IEEE. DOI: 10.1109/ICACS47775.2020.9055944

Berardi, U., & Jafarpur, P. (2020). Assessing the impact of climate change on building heating and cooling energy demand in Canada. *Renewable & Sustainable Energy Reviews*, *121*, 109681. DOI: 10.1016/j.rser.2019.109681

Bharati, P., & Pramanik, A. (2020). Deep learning techniques—R-CNN to Mask R-CNN: A survey. In Das, A., Nayak, J., Naik, B., Pati, S., & Pelusi, D. (Eds.), *Computational intelligence in pattern recognition* (*Vol. 999*, pp. 657–668). Springer., DOI: 10.1007/978-981-13-9042-5_56

Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K., & Ghayvat, H. (2021). CNN variants for computer vision: History, architecture, application, challenges and future scope. *Electronics (Basel)*, *10*(20), 2470. DOI: 10.3390/electronics10202470

Burner, J., Schwiedrzik, L., Krykunov, M., Luo, J., Boyd, P. G., & Woo, T. K. (2020). High-performing deep learning regression models for predicting low-pressure CO2 adsorption properties of metal-organic frameworks. *The Journal of Physical Chemistry. C, Nanomaterials and Interfaces*, *124*(51), 27996–28005. DOI: 10.1021/acs.jpcc.0c06334

Chen, Y., Wang, Y., Liu, X., & Huang, J. (2022). Short-term load forecasting for industrial users based on Transformer-LSTM hybrid model. In *2022 IEEE 5th International Electrical and Energy Conference (CIEEC)* (pp. 2470–2475). IEEE. DOI: 10.1109/CIEEC54735.2022.9846658

Dumas, J., Wehenkel, A., Lanaspeze, D., Cornélusse, B., & Sutera, A. (2022). A deep generative model for probabilistic energy forecasting in power systems: Normalizing flows. *Applied Energy*, *305*, 117871. DOI: 10.1016/j.apenergy.2021.117871

Fan, J., Zhang, K., Huang, Y., Zhu, Y., & Chen, B. (2023). Parallel spatio-temporal attention-based TCN for multivariate time series prediction. *Neural Computing & Applications*, *35*(18), 13109–13118. DOI: 10.1007/s00521-021-05958-z

Ferrag, M. A., & Maglaras, L. (2019). DeepCoin: A novel deep learning and blockchain-based energy exchange framework for smart grids. *IEEE Transactions on Engineering Management*, *67*(4), 1285–1297. DOI: 10.1109/TEM.2019.2922936

Glebova, E., Zare, F., Desbordes, M., & Géczi, G. (2022). COVID-19 sport transformation: New challenges and new opportunities. *Physical Culture and Sport Studies and Research*, *95*(1), 54–67. DOI: 10.2478/pcssr-2022-0011

Hafeez, G., Alimgeer, K. S., & Khan, I. (2020). Electric load forecasting based on deep learning and optimized by heuristic algorithm in smart grid. *Applied Energy*, *269*, 114915. DOI: 10.1016/j.apenergy.2020.114915

Karpov, P., Godin, G., & Tetko, I. V. (2019). A transformer model for retrosynthesis. In Tetko, I., Kůrková, V., Karpov, P., & Theis, F. (Eds.), *Artificial neural networks and machine learning – ICANN 2019: Workshop and special sessions* (*Vol. 11731*, pp. 817–830). Springer., DOI: 10.1007/978-3-030-30493-5_78

Kim, J.-a., Sung, J.-Y., & Park, S.-h. (2020). Comparison of Faster-RCNN, YOLO, and SSD for real-time vehicle type recognition. In *2020 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)* (pp. 1–4). IEEE. DOI: 10.1109/ICCE-Asia49877.2020.9277040

Kotsiopoulos, T., Sarigiannidis, P., Ioannidis, D., & Tzovaras, D. (2021). Machine learning and deep learning in smart manufacturing: The smart grid paradigm. *Computer Science Review*, *40*, 100341. DOI: 10.1016/j.cosrev.2020.100341

Lei, G., Zhuang, L., & Gan, B. (2022). Targeting ferroptosis as a vulnerability in cancer. *Nature Reviews. Cancer*, *22*(7), 381–396. DOI: 10.1038/s41568-022-00459-0 PMID: 35338310

Li, W., Wei, Y., An, D., Jiao, Y., & Wei, Q. (2022). LSTM-TCN: Dissolved oxygen prediction in aquaculture, based on combined model of long short-term memory network and temporal convolutional network. *Environmental Science and Pollution Research International*, *29*(26), 39545–39556. DOI: 10.1007/s11356-022-18914-8 PMID: 35103942

Lin, L., Fan, H., Zhang, Z., Xu, Y., & Ling, H. (2022). Swintrack: A simple and strong baseline for transformer tracking. *Advances in Neural Information Processing Systems*, *35*, 16743–16754.

Liu, Y., Zhou, D., Nie, R., Ding, Z., Guo, Y., Ruan, X., Xia, W., & Hou, R. (2022). TSE_Fuse: Two stage enhancement method using attention mechanism and feature-linking model for infrared and visible image fusion. *Digital Signal Processing*, *123*, 103387. DOI: 10.1016/j.dsp.2022.103387

Liu, Z., et al. (2020). Carbon Monitor, a near-real-time daily dataset of global CO2 emission from fossil fuel and cement production. *Scientific Data*, *7*(1), 392. DOI: 10.1038/s41597-020-00708-7 PMID: 33168822

Luo, M., Du, B., Zhang, W., Song, T., Li, K., Zhu, H., Birkin, M., & Wen, H. (2023). Fleet rebalancing for expanding shared e-Mobility systems: A multi-agent deep reinforcement learning approach. *IEEE Transactions on Intelligent Transportation Systems*, *24*(4), 3868–3881. DOI: 10.1109/TITS.2022.3233422

Lyu, T., Gu, D., Chen, P., Jiang, Y., Zhang, Z., Pang, H., Zhou, L., & Dong, Y. (2024). Optimized CNNs for rapid 3D point cloud object recognition. *arXiv preprint arXiv:2412.02855*.

Mansour, R. F., Escorcia-Gutierrez, J., Gamarra, M., Villanueva, J. A., & Leal, N. (2021). Intelligent video anomaly detection and classification using faster RCNN with deep reinforcement learning model. *Image and Vision Computing*, *112*, 104229. DOI: 10.1016/j.imavis.2021.104229

Miah, J., Mamun, M., Rahman, M. M., Mahmud, M. I., Ahmad, S., & Nasir, M. H. B. (2022). Mhfit: Mobile health data for predicting athletics fitness using machine learning models. In *2022 2nd International Seminar on Machine Learning, Optimization, and Data Science (ISMODE)* (pp. 584–589). IEEE. DOI: 10.1109/ISMODE56940.2022.10180967

Oh, G., Leblanc, D. J., & Peng, H. (2020). Vehicle energy dataset (VED), a large-scale dataset for vehicle energy consumption research. *IEEE Transactions on Intelligent Transportation Systems*, *23*(4), 3302–3312. DOI: 10.1109/TITS.2020.3035596

Pallonetto, F., De Rosa, M., Milano, F., & Finn, D. P. (2019). Demand response algorithms for smart-grid ready residential buildings using machine learning models. *Applied Energy*, *239*, 1265–1282. DOI: 10.1016/j.apenergy.2019.02.020

Pang, H., Zhou, L., Dong, Y., Chen, P., Gu, D., Lyu, T., & Zhang, H. (2024). Electronic health records-based data-driven diabetes knowledge unveiling and risk prognosis. *arXiv preprint arXiv:2412.03961*.

Pappalardo, L., Cintia, P., Rossi, A., Massucco, E., Ferragina, P., Pedreschi, D., & Giannotti, F. (2019). A public data set of spatio-temporal match events in soccer competitions. *Scientific Data*, *6*(1), 236. DOI: 10.1038/s41597-019-0247-7 PMID: 31659162

Patel, D., Shah, D., & Shah, M. (2020). The intertwine of brain and body: A quantitative analysis on how big data influences the system of sports. *Annals of Data Science*, *7*(1), 1–16. DOI: 10.1007/s40745-019-00239-y PMID: 38624463

Shen, N., Wang, T., Gan, Q., Liu, S., Wang, L., & Jin, B. (2022). Plant flavonoids: Classification, distribution, biosynthesis, and antioxidant activity. *Food Chemistry*, *383*, 132531. DOI: 10.1016/j.foodchem.2022.132531 PMID: 35413752

Shi, J., Xie, W., Huang, X., Xiao, F., Usmani, A. S., Khan, F., Yin, X., & Chen, G. (2022). Real-time natural gas release forecasting by using physics-guided deep learning probability model. *Journal of Cleaner Production*, *368*, 133201. DOI: 10.1016/j.jclepro.2022.133201

Sigmund, G., Gharasoo, M., Hüffer, T., & Hofmann, T. (2020). Deep learning neural network approach for predicting the sorption of ionizable and polar organic pollutants to a wide range of carbonaceous materials. *Environmental Science & Technology*, *54*(7), 4583–4591. DOI: 10.1021/acs.est.9b06287 PMID: 32124609

Smagulova, K., & James, A. P. (2019). A survey on LSTM memristive neural network architectures and applications. *The European Physical Journal. Special Topics*, *228*(10), 2313–2324. DOI: 10.1140/epjst/e2019-900046-x

Song, Y., Du, X., Zhang, Y., & Li, S. (2022). Two-stage segmentation network with feature aggregation and multi-level attention mechanism for multi-modality heart images. *Computerized Medical Imaging and Graphics*, *97*, 102054. DOI: 10.1016/j.compmedimag.2022.102054 PMID: 35339724

Xi, X., Zhang, C., Jia, W., & Jiang, R. (2024). Enhancing human pose estimation in sports training: Integrating spatiotemporal transformer for improved accuracy and real-time performance. *Alexandria Engineering Journal*, *109*, 144–156. DOI: 10.1016/j.aej.2024.08.072

Xiao, J., & Zhou, Z. (2020). Research progress of RNN language model. In *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)* (pp. 1285–1288) IEEE. DOI: 10.1109/ICAICA50127.2020.9182390

Xiaohua, W., Muzi, P., Lijuan, P., Min, H., Chunhua, J., & Fuji, R. (2019). Two-level attention with two-stage multi-task learning for facial emotion recognition. *Journal of Visual Communication and Image Representation*, *62*, 217–225. DOI: 10.1016/j.jvcir.2019.05.009

Zhang, L., Xu, C., Gao, Y., Han, Y., Du, X., & Tian, Z. (2020). Improved Dota2 lineup recommendation model based on a bidirectional LSTM. *Tsinghua Science and Technology*, *25*(6), 712–720. DOI: 10.26599/TST.2019.9010065

Zhao, Y., & Li, H. (2023). Understanding municipal solid waste production and diversion factors utilizing deep-learning methods. *Utilities Policy*, *83*, 101612. DOI: 10.1016/j.jup.2023.101612

Zhu, H., Goh, H. H., Zhang, D., Ahmad, T., Liu, H., Wang, S., Li, S., Liu, T., Dai, H., & Wu, T. (2022). Key technologies for smart energy systems: Recent developments, challenges, and research opportunities in the context of carbon neutrality. *Journal of Cleaner Production*, *331*, 129809. DOI: 10.1016/j.jclepro.2021.129809