


Review

# Deep and Reinforcement Learning in Virtual Synchronous Generator: A Comprehensive Review

Xiaoke Ding<sup>1</sup> and Junwei Cao<sup>2,\*</sup> 

<sup>1</sup> Department of Automation, Tsinghua University, Beijing 100084, China; dxk21@mails.tsinghua.edu.cn

<sup>2</sup> Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China

\* Correspondence: jcao@tsinghua.edu.cn

**Abstract:** The virtual synchronous generator (VSG) is an important concept and primary control method in modern power systems. The penetration of power-electronics-based distributed generators in the power grid provides uncertainty and reduces the inertia of the system, thus increasing the risk of instability when disturbance occurs. The VSG produces virtual inertia by introducing the dynamic characteristics of the synchronous generator, which provides inertia and becomes a grid-forming control method. The disadvantages of the VSG are that there are many parameters to be adjusted and its operation process is complicated. However, with the rapid development of artificial intelligence (AI) technology, the powerful adaptive learning capability of AI algorithms provides potential solutions to this issue. Two research hotspots are deep learning (DL) and reinforcement learning (RL). This paper presents a comprehensive review of these two techniques combined with VSG control in the energy internet (EI). Firstly, the basic principle and classification of the VSG are introduced. Next, the development of DL and RL algorithms is briefly reviewed. Then, recent research on VSG control based on DL and RL algorithms are summarized. Finally, some main challenges and study trends are discussed.

**Keywords:** virtual synchronous generator; artificial intelligence; deep learning; reinforcement learning



**Citation:** Ding, X.; Cao, J. Deep and Reinforcement Learning in Virtual Synchronous Generator: A Comprehensive Review. *Energies* **2024**, *17*, 2620. <https://doi.org/10.3390/en17112620>

Academic Editor: Armando Pires

Received: 6 April 2024

Revised: 26 May 2024

Accepted: 26 May 2024

Published: 29 May 2024



**Copyright:** © 2024 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/>).

## 1. Introduction

At present, the power system is experiencing ongoing transformation and reconstruction in order to be more flexible, sustainable, distributed, environmentally friendly, and intelligent. Traditional large thermal power plants are gradually replaced by distributed generators (DGs) with renewable energy sources (RESs), such as wind and solar. This means a shift from centralized power generation and long-distance power transportation to distributed generation and regional energy autonomy. The energy storage system (ESS), which allows power to flow in both directions, is applied to deal with the volatility of RESs and the randomness of local loads. The concept of the energy internet (EI) was first proposed by American economist Jeremy Rifkin in his book “The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World” [1]. The principle of the EI combines the technologies of internet, power electronics, large data processing, intelligent monitoring, and management with power systems [2]. The use of power electronic converters to connect the RES generators and ESS to the power grid brings a low inertia to the system and the operation conditions are more complicated. Inertia plays an important role in system stability since a large inertia means a strong ability to stay in place when disturbance occurs. To deal with this problem of the lack of inertia, the research pays much attention to the control strategies for inverter-based distributed generators (IBDGs) to participate in frequency and voltage regulation [3–5].

To provide inertia and improve system stability, virtual synchronous generator (VSG) control technology has been proposed and widely studied. As the name VSG suggests,

this control strategy is brought up by imitating the behavior of traditional synchronous generators (SGs) [6].

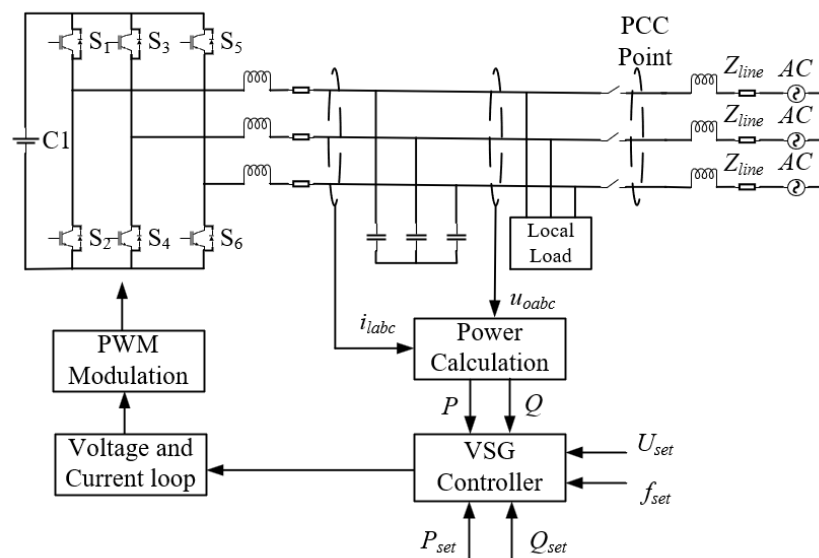
In order to better apply the VSG in the energy internet, many works in the literature take different methods and optimize different aspects for the existing VSG framework. As the current artificial intelligence (AI) methods are widely used in various traditional fields, they bring new breakthroughs to these traditional industries. This paper will focus on the most popular deep learning (DL) and reinforcement learning (RL) methods in the current AI methods combined with VSG application. To the best of our knowledge, there are no review papers on the subject combining VSG control with DL and RL methods. This paper will fill the vacancy in this field.

The structure of this paper is as follows:

- Section 2 provides an introduction of VSG principles. The main classifications of the VSG are introduced in detail;
- Section 3 gives an overall review of the development of deep learning and reinforcement learning;
- Section 4 focuses on the dynamic process of a single VSG that is optimized by DL and RL algorithms;
- Section 5 reviews the DL and RL algorithms used in multi-VSG systems;
- Section 6 concludes this paper and provides some future research trends according to all the research reviewed in this paper.

## 2. Principle and Classification of VSG

In general, a simplified VSG control system is shown in Figure 1, where  $L_r$  and  $C_r$  are, respectively, the filter inductance and filter capacitance of the inverter,  $Z_{line}$  is the line impedance of the inverter output to the AC bus, and  $P$  and  $Q$  are, respectively, the active power and reactive power output. The power supply on the DC side represents distributed power supplies such as the photovoltaic and energy storage systems. And the AC side can either operate in the off-grid mode to support local loads, or provide inertia to the grid in the grid-connected mode.



**Figure 1.** Simplified VSG system.

In this kind of imitation, the transient process of the SG is described mathematically, and then the mathematical equations are introduced into the control process. Since the transient process is complex and it is unnecessary to fully integrate the characteristics of the SG into the control process, various models with slight differences have been proposed. These models and principles are discussed in this section.

### 2.1. Virtual Synchronous Machine (VISMA)

The VISMA is the first VSG algorithm proposed by Hans-Peter Beck in [7]. This scheme simulates the rotor motion equation and stator electrical equation of the SG, and obtains the reference value of the stator current or the reference value of the motor port voltage according to the volt–ampere characteristic of the stator electrical equation. The external characteristics of the synchronous generator are simulated by controlling the output current or output voltage of the inverter.

The swing equation for the rotor can be expressed as Equation (1):

$$T_m - T_e = J \frac{d\omega}{dt} + k_d f(s) \frac{d\omega}{dt} \tag{1}$$

where  $k_d$  is the mechanical damping coefficient and  $f(s)$  is the phase compensation term [8].

The stator electrical equation can be expressed as Equation (2):

$$e = u + R_s i_L + L_s \frac{di_L}{dt} \tag{2}$$

where  $e = [e_a, e_b, e_c]$  is the three-phase stator-induced electromotive force,  $u = [u_a, u_b, u_c]$  is the three-phase voltage at the motor port,  $i_L = [i_{La}, i_{Lb}, i_{Lc}]$  is the three-phase stator current, and  $R_s$  and  $L_s$  are stator resistance and inductance values, respectively.

According to the different implementations of the Equation (1), the VISMA can be divided into the current-control type [7] and the voltage-control type [8]. The current-type VISMA scheme is based on Equation (1), obtaining the reference value of the three-phase stator current  $i_{Lref} = [i_{Laref}, i_{Lbref}, i_{Lcref}]$ :

$$i_{Lref} = \frac{1}{L_s} \int (e - u - R_s i_{Lref}) dt \tag{3}$$

While the voltage-type VISMA scheme obtains the reference value of the three-phase motor port voltage  $u_{ref} = [u_{aref}, u_{bref}, u_{cref}]$ :

$$u_{ref} = e - R_s i_L - L_s \frac{di_L}{dt} \tag{4}$$

Figure 2 shows the block diagrams of these two types of VISMAs.

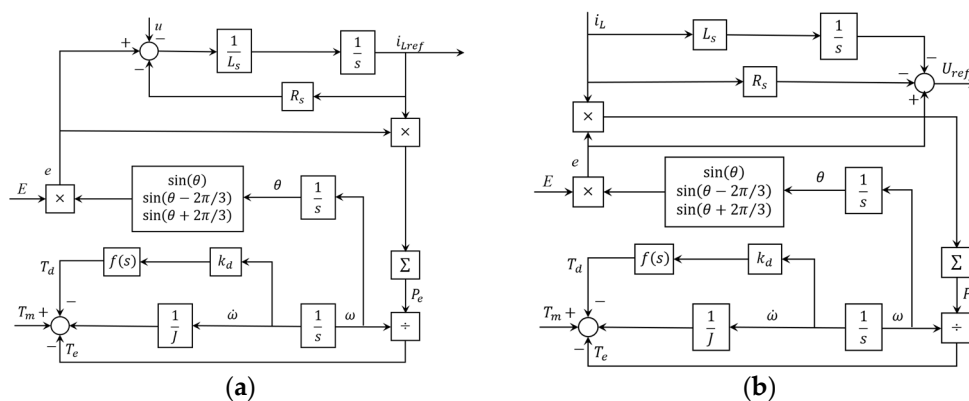


Figure 2. Diagrams of VISMA: (a) current-control type; and (b) voltage-control type.

### 2.2. Synchronverter

The synchronverter [9] simulates the operation mechanism of the synchronous generator more precisely, which not only considers the rotor motion equation of the SG in the control process, but also mimics the electromechanical coupling characteristics of the synchronous generator. The synchronverter fully takes into account the electromechanical

and electromagnetic transient characteristics of the SG and is obtained according to the following physical and mathematical models:

$$e = \omega M_f i_f \left[ \sin \theta, \sin \left( \theta - \frac{2}{3\pi} \right), \sin \left( \theta + \frac{2}{3\pi} \right) \right] \tag{5}$$

$$Q = -\omega M_f i_f \left[ i_{La} \cos \theta + i_{Lb} \cos \left( \theta - \frac{2}{3\pi} \right) + i_{Lc} \cos \left( \theta + \frac{2}{3\pi} \right) \right] \tag{6}$$

$$T_e = M_f i_f \left[ i_{La} \sin \theta + i_{Lb} \sin \left( \theta - \frac{2}{3\pi} \right) + i_{Lc} \sin \left( \theta + \frac{2}{3\pi} \right) \right] \tag{7}$$

where  $M_f$  is the mutual inductance between excitation windings and stator windings, and  $i_f$  is the exciting current.

The synchronverter strategy can be applied both in the power-side grid-connected inverter control, as shown in Figure 3a, and the load-side grid-connected rectifier control, as shown in Figure 3b. This makes the DGs and loads have the same operating mechanism as the synchronous units, allowing them to participate in the operation and management of the EI. When the power grid, DGs, or load disturbance occurs, the DGs and loads can be adjusted bi-directionally to improve the ability to resist external disturbance.

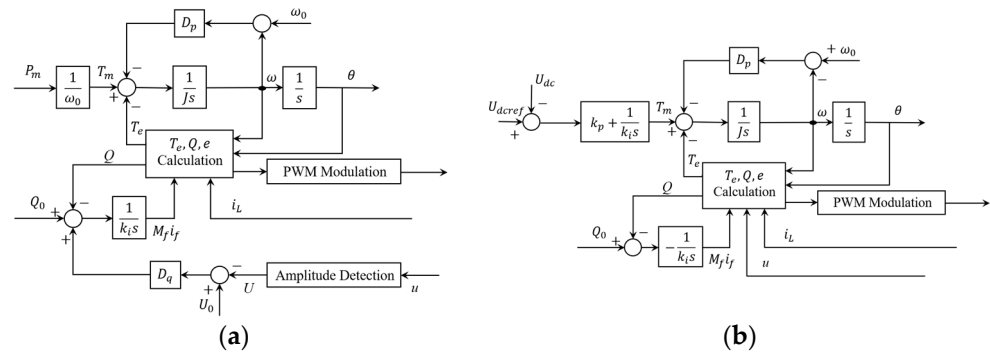


Figure 3. Synchronverter: (a) power side; and (b) load side.

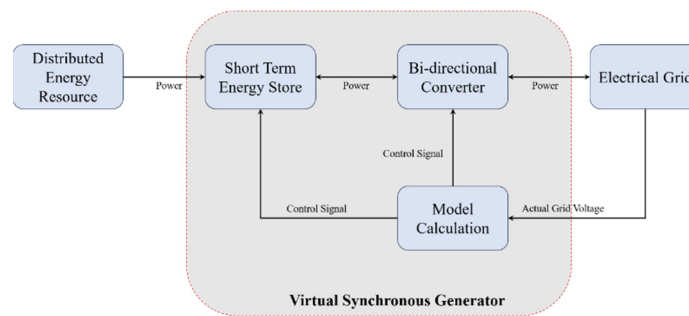
The synchronverter applied on the power side and the load side as mentioned above can be interpreted as a virtual synchronous generator and a virtual synchronous motor (VSM). The VSG type calculates the reactive power  $Q$  and electromagnetic torque  $T_e$  according to Equations (6) and (7). It realizes the function of primary frequency regulation and damping with the damping coefficient  $D$ , and simulates the rotor inertia of the synchronous motor by the rotor motion equation. In order to realize the reactive power-voltage droop control, the difference between the grid voltage rating and the actual voltage is multiplied by a factor  $D_q$  and added to the reactive power rating  $Q_0$ .

### 2.3. Other Virtual Synchronous Models

From 2008, several different VSG models are proposed worldwide. In 2008, K. VISSCHER et al. proposed the concept of the VSG in the European VSYNC joint project [10]. The first-order sag characteristic curve and dynamic mechanical equation of the generator are simulated, and the moment of inertia and damping characteristics are simulated by adding short-term energy relief. The basic principle is shown in Figure 4.

Moreover, the Institute of Electrical Power Engineering (IEPE) at Clausthal University of Technology in Germany, the VSG research team at Kawasaki Heavy Industries (KHIs), and the ISE laboratory at Osaka University in Japan have proposed several different models with which to provide inertia.

In the remaining part of this paper, the models mentioned above are collectively referred to as VSGs and no specific distinction is made.



**Figure 4.** Principle of VSG proposed in VSYNC Project.

### 3. Deep Learning and Reinforcement Learning

DL and RL are both popular branches of machine learning (ML). There are some differences in the problems they focus on and the tasks they solve.

#### 3.1. Deep Learning

Deep learning is a subset of machine learning that emphasizes the use of deep neural networks (deep-learning models) to learn and represent data [11]. DL models are able to learn complex features and representations from large amounts of data, enabling them to achieve remarkable achievements in the fields of vision, speech, natural language processing, etc.

In this section, some mainstream types of DL algorithms are briefly introduced. DL typically refers to a deep neural network which contains multiple hidden layers, as well as input and output layers. Generally, DL algorithms deal with tasks such as the following: (i) supervised: the input data are labeled; (ii) unsupervised: the input data are not labeled or without a known result; and (iii) semi-supervised: the input data are a mixture of labeled and unlabeled examples. For different learning tasks, neural networks with different structures are used.

##### 3.1.1. Multi-Layer Perceptron (MLP)

The multi-layer perceptron is a supervised learning approach and a fundamental structure of DL. There is an input layer to receive data, an output layer to make a decision, and several hidden layers to compute and process data in an MLP. The front and back layers in the MLP are typically fully connected.

##### 3.1.2. Convolutional Neural Network (CNN)

The convolutional neural network was first proposed by LeCun et al. in 1998 [12]. The CNN is suitable for processing spatial data. Its design ideas are inspired by visual neuroscience. A CNN mainly consists of a convolutional layer and a pooling layer. The convolution layer can maintain the spatial continuity of the data structure and extract the local features of the data. The pooling layer can be max-pooling or mean-pooling, which can reduce the dimension of the middle hidden layer, reduce the calculation of the next layer, and provide rotation invariance. CNN is widely used in computer vision; the ImageNet large scale visual recognition competition (ILSVRC) has greatly promoted the development of convolutional neural networks, and many classic neural networks were born in this competition.

AlexNet [13], the winner in 2012, brought about a range of technologies that are still widely used today. It adds training data by means of translation, flipping, and capturing part of the image, uses dropout to prevent overfitting, trains the model with batch gradient descent with momentum and weight decay, and uses a Rectified Linear Unit (ReLU) as an activation function to shorten the training time. Zeiler and Fergus used deconvolution networks to visualize the CNN as a way to understand the role of each layer of the CNN, which led to the 2013 winner, ZF Net [14]. VGGNet, proposed by Simonyan et al. in 2014, showed that deeper networks can significantly improve the effectiveness of deep

learning [15]. The champion of the same year, GoogleNet, placed the convolution layer and the pooling layer in parallel, breaking the pattern of the convolution-layer–pooling-layer stacking for the first time [16]. Deep Residual Network (ResNet) [17] is the ILSVRC 2015 champion. The same network won the three tasks of picture classification, object location, and object detection. ResNet directly adds a linear connection path between two or more layers to form a residual module, which effectively solves the problem of gradient disappearance in deep networks. R-CNN [18] and its many improved methods [19–21] are the mainstream methods applied to object detection.

### 3.1.3. Recurrent Neural Networks

Recurrent neural networks (RNNs) take the output of the hidden layer of the previous moment as the input of the hidden layer of this moment, and can use the information of the previous moment [22]. This property of memory is well-suited for processing time-series data. The basic RNN principle can be illustrated as follows: RNN shares the weight at each time, and, when using information from a long time in the past, there will be a gradient disappearance or gradient explosion problem [23].

The long short-term memory (LSTM) network replaces neurons in the RNN with LSTM units, adding input gates, output gates, and forget gates to input, output, and forget information to control how much information is allowed to pass through [24]. The LSTM model can remember things that need to be remembered for a long time and forget unimportant ones, which can alleviate the problems of gradient disappearance and gradient explosion, and have a better performance than RNN on longer sequences. The gated recurrent unit (GRU) is a lightweight variant of the LSTM model with only two doors: an update door and a reset door [25]. The update gate determines how much past information is retained and how much information is entered from the input layer; The reset door is similar to the forget door in LSTM. With less connection and fewer parameters, the training is faster compared with LSTM. Some other improvements like the hierarchical RNN [26], bidirectional RNN [27] and multi-dimensional RNN [28] have been proposed to deal with different problems arising in the RNN.

### 3.1.4. Attention Mechanism

When people observe things through the visual system, the human eye focuses on certain parts (of interest). Similarly, the attention mechanism enables deep-learning models to assign different weights to different input data, giving more attention to more important data, which is realized by incorporating mechanisms to focus on the moving, scaling, and rotating of the areas into the network [29]. The attention mechanism can be combined with the CNN [30], RNN [31], and LSTM [32].

The classification of supervised DL mentioned above can be classified as shown in Figure 5.

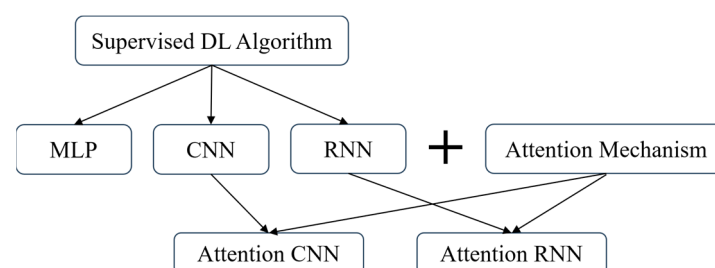


Figure 5. Classification of DL algorithms.

## 3.2. Reinforcement Learning

Reinforcement learning is a machine-learning method in which agents interact with the environment to learn optimal behavioral strategies. Agents adjust their behavioral strategies based on feedback from the environment (reward or punishment). RL focuses on

developing the best decision-making strategy by maximizing long-term rewards through trial and error and learning. It is widely used in decision problems, control systems, and automation.

The development of reinforcement learning can be traced back to the 1950s and 1960s, and the founding work can be traced to Richard Bellman's dynamic programming theory. Bellman first proposed the Bellman equation for optimal control problems, which laid the foundation for the theory of reinforcement learning [33]. In the 1960s, Ronald Howard et al. introduced the Monte Carlo method into optimal control problems [34]. Later, Ronald Howard and Andrey Kolmogorov et al. formally described the Markov decision process (MDP), which became an important mathematical model in reinforcement learning. The MDP provides a theoretical framework for describing the interaction between agents and the environment and making decisions. Christopher Watkins proposed the Q-learning algorithm in [35] in 1989. RL combined with deep learning, which is referred to as deep reinforcement learning (DRL), has developed rapidly in the past decade. New algorithms have been constantly proposed and the application fields have become more and more extensive. To be more specific, RL in this paper mainly refers to the deep reinforcement method.

There are two main components in an RL algorithm. The agent can interact with the environment and obtain a reward from it. The agent explores the action space and finds out which action under the current state can bring the greatest cumulative reward [36]. Different from the immediate rewards of supervised and unsupervised learning, the rewards in RL are long-term and emphasize the process of accumulation.

A reinforcement-learning problem can be described as a Markov decision process (MDP), which consists of a set of interacting objects, the agent, and environment [37]. A reinforcement-learning agent can make decisions by sensing the state of the external environment, take actions, and adjust decisions through the feedback of the environment. The environment refers to all things outside the agent in the MDP model, the state of whom is changed by the actions of the agent. The changes can be fully or partially perceived by the agent. A positive change may give a reward to the agent while a negative one can bring a punishment, causing the agent to adjust accordingly.

MDP can be described as a tuple  $\langle S, A, P, R \rangle$ , where:

- $S = \{s_1, s_2, \dots, s_t\}$  stands for the state space, which can be either discrete or continuous. It refers to the state of the environment observed by the agent.
- $A = \{a_1, a_2, \dots, a_t\}$  stands for the action space taken by the agent, which can be either discrete or continuous.  $a_t$  refers to the action produced by the agent decision at each time step  $t$ . After each action, the environment will enter the next state.
- $P$  refers to the transfer probability function, which represents the possibility that the system moves to the next state after the agent takes a certain action in the current state.
- $R$  refers to the reward function. By interacting with the environment, the agent will receive a reward for its action. A positive reward indicates the action is effective, while a negative reward, also known as a punishment, indicates a wrong action. The goal of an agent is to maximize the expected future rewards by optimizing the policy.

Besides the elements above, a policy  $u(a_t|s_t)$  is a function that maps the state  $s_t$  to the action  $a_t$ . The ultimate goal of an RL algorithm is to find out a certain policy to maximize the benefits. The illustration of an MDP is shown in Figure 6. Every time the agent takes an action  $a_t$ , the environment generates a reward  $r_t$  based on the current state  $s_t$ , and then moves to the next state  $s_{t+1}$ .

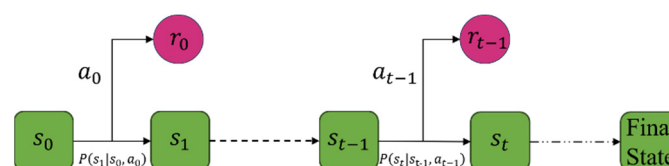


Figure 6. MDP illustration.

For an MDP, the state-value function  $V_\pi(s)$  and the action-value function  $Q_\pi(s, a)$  are used to describe how acceptable the action taken in a certain state is. The state-value function  $V_\pi(s)$  can be defined as:

$$V_\pi(s) = \mathbb{E}_\pi[R_t | S_t = s] \quad (8)$$

which means the expected reward value by following the policy  $\pi$  in condition  $s$ .  $R_t$  is the estimated value of the cumulative future reward defined as:

$$R_t = r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \dots = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \quad (9)$$

where  $\gamma$  is a discount factor and  $\gamma \in [0, 1)$ .

The action-value function  $Q_\pi(s, a)$  can be defined as:

$$Q_\pi(s, a) = \mathbb{E}_\pi[R_t | S_t = s, a_t = a] \quad (10)$$

which means the expected long-term benefits by taking the action  $a_t$  in the state  $s_t$ .

The optimal policy  $\pi^*$  refers to the policy that receives the highest cumulative reward. And the optimal state-value function and action-value function can be defined as:

$$V^*(s) = \max_{\pi} V_\pi(s) \quad (11)$$

$$Q^*(s, a) = \max_{\pi} Q_\pi(s, a) \quad (12)$$

To find out the optimal policy, RL algorithms can be divided into two categories: value-based approaches and policy-based approaches.

### 3.2.1. Value-Based Approaches

The deep Q-network (DQN) is a representative-value-based approach that marks the beginning of the flourishing of deep reinforcement learning. Developed by DeepMind, a DQN algorithm has outperformed human players in an Atari 2600 game, demonstrating the potential of deep learning in reinforcement learning [38]. In the DQN, a DNN is used to replace the tabular value in classic Q-learning. Moreover, many of the contributions of the DQN are groundbreaking. First, an experienced replay mechanism based on a replay buffer is developed to store transition tuples. Then, a target network is introduced in the process. The target network  $Q(s, a|\theta')$  has the same structure with the online network  $Q(s, a|\theta)$  but with different parameters. The parameters  $\theta'$  in the target network update slowly to overcome the overfitting problem. It is widely confirmed that learning with a target network performs better in many tasks [39]. Furthermore, an  $\epsilon$ -greedy strategy is applied to encourage exploration to prevent the agent from acting within the explored space.

The Double DQN (DDQN) [40], dueling DQN [41], and Prioritized DQN [42] are improved approaches based on the DQN to solve the problem of action-value function estimation. In the DDQN, two neural networks are applied to select actions and calculate the target Q value, respectively, which helps to solve the problem of overestimation in the DQN, and improves the stability and performance of training, while, in the dueling DQN, the Q-value function is divided into two parts: the value function of the state and the advantage function of the action. The dueling DQN can learn the value information of states and the advantage information of actions, thus achieving a faster convergence. The prioritized DQN innovatively takes the temporal difference (TD) as an important consideration to ensure that important experiences can be replayed and learned first, which greatly improves the learning efficiency of the DQN.

### 3.2.2. Policy-Based Approaches

The value-based approaches output the value of all actions and select the action with the highest value, which takes the same action under the same circumstance. The policy-based approaches parameterize the strategy, take the weight parameter of the neural



network as the parameter of the value function, and directly output the probability of various actions to be taken in the next step by analyzing the state. The actions are taken according to the probability. Each action has a corresponding probability to be selected.

Compared with value-based approaches, policy-based approaches are simpler to realize and have a better convergence. However, the size of the learning step for policy-based approaches is more difficult to determine. The trust region policy optimization (TRPO) algorithm introduces Kullback–Leibler divergence (KLD) to represent the gap between old and new strategies, and can finally solve the maximum step length in a confidence region that can make the strategy monotonically improve, reducing the fluctuation during training [43]. The proximal policy optimization (PPO) algorithm innovatively uses a new loss function, Clip, instead of KLD to solve the problems of the large computational load and complex implementation of TRPO [44]. The authors of [45] proposed a deterministic policy gradient (DPG) algorithm theoretically and provide the calculation formula and parameter-updating method. As suggested in this name, the deterministic method greatly improves the training speed compared with the stochastic strategy gradient method.

Actor–Critic is a policy-based approach that blends TD learning and a policy gradient. The actor refers to the policy function  $\pi(a|s)$  and the critic refers to the value function  $V_{\pi}(s)$ . The deep deterministic policy gradient (DDPG) algorithm improves the DPG by using this Actor–Critic architecture to replace the linear function for the value function prediction through the neural network [46]. Since the action generated by the DPG in the same state is determined, the DDPG combined the advantages of the DPG and the DQN. Using the Actor–Critic Framework, the Actor to explore and the Critic to revise, it ensures the ability to explore the action space and improves the computational efficiency. Based on the Actor–Critic framework, various algorithms have been proposed, such as the Advantage Actor–Critic (A2C), Asynchronous Advantage Actor–Critic (A3C) [47], Soft Actor–Critic (SAC) [48], and twin delayed deep deterministic policy gradient (TD3) algorithm [49].

The RL-based approaches can be classified as shown in Figure 7.

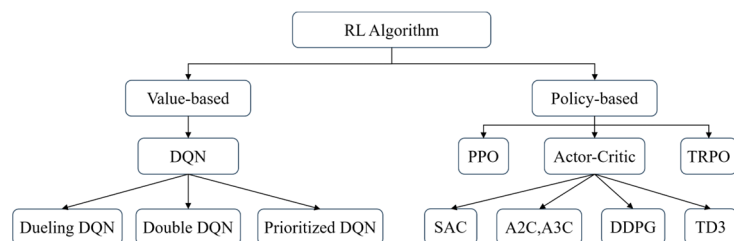


Figure 7. Classification of RL algorithms.

#### 4. DL and RL in Single VSG

The aim of the VSG is to mimic the inertia and damping characteristic of SGs. Apart from that, one of the advantages of the VSG is its flexibility. The parameters of SGs are fixed since the physical structure and state are fixed. But the parameters in the VSG are flexible towards change since the control algorithm is just a program running on a controller. To make better use of the flexibility of the VSG, various parameter adaptive control methods have been widely studied.

The mode-based method depends on the accurate system model to develop a certain adaptive law for tuning. In [50], an adaptive virtual inertia control method is proposed to improve the dynamic performance of the system. In [51], both virtual inertia and damping are taken into consideration. The effect of control parameters on the output power and frequency have been intuitively studied in [52] and a strategy with the features of a short response time and small overshoot is proposed.

With the rapid development of AI technology, its powerful ability to learn data relationships causes it to have a good application prospect in parameter adjustment.

DL algorithms are introduced to deal with this issue. In [53], taking the lifetime consumption into consideration, a double-ANN-based method is proposed to design

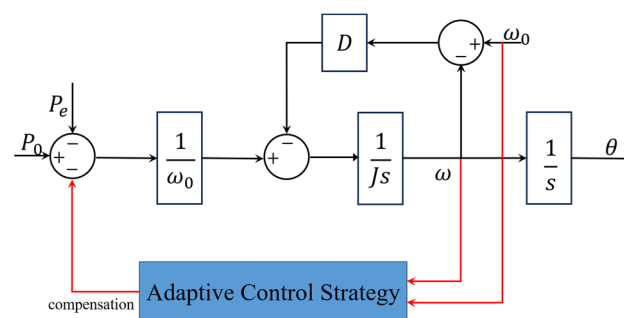
the virtual inertia with stability and reliability. A deep convolutional neural network (DCNN) is designed in [54] to adjust the virtual inertia  $J$  and damping  $D$  online. The input data for the net are the frequency deviation and the rate of change of frequency. The convolution layer, pooling layer, and fully connected layer are involved in the network. The results show a better performance over the fixed-parameter method in active power and frequency response.

RL algorithms are also widely used in this field. Q-learning is first used in VSG parameter tuning in [55]. Even though the structure design is relatively simple (the virtual inertia  $J$  is the only one-dimensional action output, and the reward function only considers the deviation of angular frequency), the method can still effectively improve the dynamic frequency and active power response of the VSG. In [56], a DQN algorithm is proposed, which has a continuous state set by replacing the Q-table in Q-learning with a neural network. The virtual inertia  $J$  and damping  $D$  are selected as actions with discrete action spaces. The deviation of active power, reactive power, and frequency are all considered in the reward function. Reference [57] applies a decentralized DPG algorithm in order to obtain continuous control operations with continuous state observations. And the derivative of angular frequency, that is, the rate of change of frequency (RoCoF), is taken into consideration in order to obtain relatively slow and smooth electromechanical dynamics. However, when applied in a new system, the results of the proposed methods show no significant advantage over the droop controller. In [58], a physics-informed neural network combined with an SAC algorithm is proposed to adjust the virtual inertia  $J$  and virtual damping  $D$  of a VSG. To decrease the risk of grid instability, physical constraints are taken into consideration. The proposed method is tested in the IEEE-39 bus simulation system. Reference [59] uses an RL algorithm based on Actor-Critic to adjust the virtual inertia. In [60], virtual inertia is adjusted by a DDPG agent to achieve a better dynamic frequency response. A DDPG algorithm is proposed in [61] to adjust  $J$  and  $D$  adaptively. The author of [62] argues that the DDPG algorithm may suffer from overestimation bias and proposed a TD3 algorithm. Moreover, the settling time is involved in the reward function to ensure a fast response. Considering different reward functions and various working conditions, the results of the proposed TD3 method are fully compared with some other control strategies both in Simulink and a real-time digital simulator (RTDS). In [63], the VSG control is applied in a modular multilevel converter (MMC) structure and a TD3 algorithm is used to adjust the virtual inertia  $J$ , damping  $D$ , and the voltage reference  $E$  according to the deviation of active power, reactive power, and frequency. The TD3 algorithm is also used to adjust  $J$  and  $D$  in [64], in which the authors make a comparative analysis between TD3 and other two RL algorithms, DDPG and SAC. Different methods are compared both in control performance and in training performance. In [65], an adaptive dynamic programming (ADP) algorithm is used to adjust the virtual inertia of a VSG and the angular deviation, frequency deviation, and RoCoF are taken into consideration. Compared with other research, Ref. [65] provides a bounded threshold of inertia to facilitate practical applications.

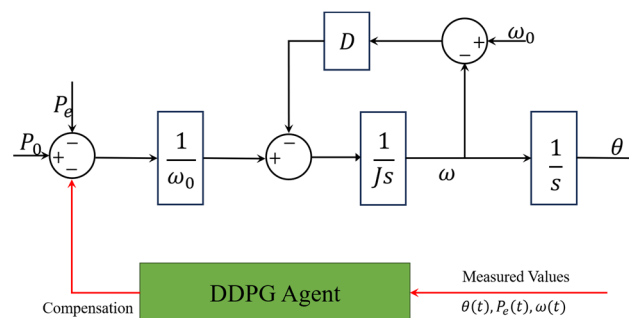
Besides virtual inertia and damping, some researchers use DL or RL to adjust other parameters in the VSG. In [66], a DDPG algorithm is used to adjust the active power reference value. Compared with the PI controller, the DDPG method achieves a better performance in the grid-connected mode and realizes the frequency restoration in the islanded mode. Reference [67] applies a safe RL method with which to adjust the active charging power reference of a battery energy storage system (BESS). The transient frequency regulation process is improved compared with the DDPG-based method and the SAC-based method. In [68], different grid conditions like an unbalanced grid, voltage drop, and a weak grid are all included. The ADP algorithm is used to realize the secondary control. The proposed method is tested both in simulation and in a lab. In [69], a model-dependent heuristic dynamic programming (MDHDP) algorithm is introduced. Receiving the power, frequency, and angle information of the system, the actor network outputs the peak value of the inverter's voltage. And the simulation results show the effectiveness of the proposed

method. The authors of [69] further carry out the examination on a test bed and the results are shown in [70].

Some research provides other solutions with which to improve the dynamic performance of the VSG. Instead of adjusting some core parameters in a reasonable range, another natural idea is to compensate the variables that the system needs to control. The traditional active power loop of the VSG is shown in Figure 8, while, in [37], a power compensation term is introduced to this loop, as shown in Figure 9. The action output of the DDPG agent is the power compensation to the input of the active power loop. With fixed parameters  $J$  and  $D$ , the active power and frequency responses of the system are also improved. In [71], the author replaces the VSG controller with a DNN with five layers. The data generated by a traditional VSG control system under multiple working conditions are selected as the training data. The inputs of the network include the output voltages and currents of the inverters, loads, angular frequency, the required phase voltage, and frequency. The outputs are the duty ratio under the dq-frame and the angular frequency, which are sent to the PWM module after the Parker inverse transform calculation to control the inverters. In [72], the LSTM algorithm is applied to mimic the relationship between the inputs and outputs of a VSG control model when a phase-to-phase short circuit fault occurs. And the same idea is applied in [73] by combining the restricted Boltzmann machine (RBM) features and a DNN. With fewer inputs, the proposed controller can suppress the power oscillation and provide a stable frequency. The effectiveness is verified on an experimental platform in a laboratory.



**Figure 8.** Traditional adaptive control for VSG frequency regulation.



**Figure 9.** Method proposed in [71] for frequency regulation.

The authors of [74] focus on the power-coupling problem of VSG. A new reactive-voltage loop is designed, which is symmetric to the active-frequency loop. By using an ADP algorithm, the parameters of the two loops above are adjusted. The constraints of PQ decoupling are also relaxed. The results performed on a test rig shows the effectiveness of the proposed method.

The references mentioned in this section and the methods used are shown in Table 1.

Table 1. DL and RL in single VSG control.

References	Learning Algorithm		Power	Objective Frequency	Voltage	Adjust Parameters	Method Replace VSG Module
	DL	RL					
[53]	✓					✓	
[54]	✓		✓	✓		✓	
[55]		✓	✓	✓		✓	
[56]		✓	✓	✓		✓	
[57]		✓	✓	✓		✓	
[58]		✓	✓	✓		✓	
[59]		✓	✓	✓		✓	
[60]		✓	✓	✓		✓	
[61]		✓	✓	✓		✓	
[62]		✓	✓	✓		✓	
[63]		✓	✓	✓	✓	✓	
[65]		✓	✓	✓		✓	
[66]		✓	✓	✓		✓	
[67]		✓	✓	✓		✓	
[69]		✓		✓	✓		✓
[71]	✓			✓	✓		✓
[72]	✓			✓	✓		✓
[73]	✓		✓	✓			✓
[74]		✓	✓			✓	

## 5. DL and RL in Multi-VSG System

Power systems with regional energy internet or microgrids have the ability of local autonomy. The local power supply can be fully absorbed by local loads or ESSs and scheduled by a local governor, which requires that all parts of the microgrids work together. The co-ordinated control of a multi-VSG system consists of system transient stability prediction, power sharing, the suppression of current circulation, and power oscillation.

### 5.1. Power Sharing, Suppression of Current Circulation, and Power Oscillation

It can be drawn from the discussion above that the operation mechanism of the VSG is relatively more complex than other control methods. As a result, the dynamics of several paralleled VSGs can interact with each other. It is shown in [75,76] that the placement of virtual inertia and damping, “where” to place, rather than the total inertia and damping in a system, has a more significant influence on the efficiency and stability. Since then, more and more researchers have directed their attention towards the co-ordinated operation of a distributed system.

Some learning methods like particle swarm optimization (PSO) in [77,78] are introduced to attain optimal multi-VSG parameters on a detailed model. But these parameter optimization methods are offline and based on fixed working conditions. Some model-based methods in [79–83], realized by the dynamic model of the VSG, can attain optimal parameters and achieve a good performance. However, an accurate mathematical model of the multi-VSG system is hard to obtain and the working conditions change over time. Therefore, the model-free methods attract more attention in application.

It is analyzed in [84] that all VSGs output the same frequency when the virtual inertia, droop coefficient, and the equivalent disturbance are proportional. By adjusting the parameters of each VSG in real time, the oscillation can be suppressed. According to this theory, the author raises an SAC-based model-free method to modify the inertia and droop coefficient, which shows a good robustness under communication interruption and time delay.

In [85], an SAC algorithm is used to adjust the droop coefficients and the DC voltage reference values in a system with four AC/DC converters. Compared with the traditional fixed parameters control method, the SAC method performs better in frequency support when the transmission load is heavy.

In [86], to design a co-ordinated wide-area damping control method, two control schemes are proposed, one based on the DNN and the other based on the DDPG. The different two agents can both track the system dynamics and take actions to improve safety

and reliability under various conditions. In [87], an ANN-based virtual inertia control system is introduced to support the frequency of a multi-area microgrid system. And the multi-agent RL method is used to find out the weights of each ANN. Compared with the PID method in a three-area interconnected system, the proposed method can reduce oscillations at high frequencies.

Reference [88] applies a TD3 algorithm to adjust the reactive power reference for each VSG to minimize the reactive-power-sharing error and ensure voltage stability. In [89], the multi-agent distributed frequency control strategy based on Q-learning is applied to realize frequency restoration and active power sharing among the four VSGs. The control actions are the active power compensation values for VSGs.

### 5.2. Transient Stability Prediction and Improvement

With the high penetration of power electronic devices, the nonlinear characteristics make it difficult to analyze the transient stability by traditional analysis methods. The development of AI technology brings solutions to this complicated process. Some ML methods like the Decision Tree (DT) and Support Vector Machine (SVM) are first raised in this field. In [90], phasor measurements and the DT are used to assess the dynamic security of a large-scale power system with a 2100-bus, 2600-line, and 240-generator model, which reaches an accuracy of 97%.

Afterwards, DL methods are introduced to this field. An online monitoring system (OMS) is proposed in [91] using a convolutional neural network (CNN). The OMS monitors the voltage profiles continuously, makes predictions on stability, and provides the most impacted generators. In [92], a deep feed-forward neural network (DFFNN) method is applied for transient stability prediction in a 3-VSG system. Compared with shallow ML methods, the DFFNN method realizes a more rapid prediction with 0.18 ms and a higher accuracy of 97.2% under complex working conditions. In [93], the author raises a message-passing graph neural network (MPGNN) to make it possible to use the same algorithm under different topologies of the system without training again, which improves the generalization ability compared with [92]. Besides accuracy, some of the research focuses on efficiency. Only using the virtual power angle data of each VSG, the long-short term memory (LSTM) neural network proposed in [94] takes 0.02 s to predict the virtual power angle of the next 0.25 s. To further reduce the amount of computation and the network size, a variable kernel-temporal convolutional network (VK-TCN) method is proposed in [95], which takes a minimum of 43 s to train. By analyzing the mechanism of voltage instability in the island microgrid, it narrows the range of eigenvalues and reduces trial and error in adjusting the hyperparameters of the network. Another control method to improve transient stability is called preventive control. According to [96], preventive control changes the operating point at a small cost by adjusting the power output of generators or the load configuration before a large disturbance to meet the requirement of transient stability. The solution to the preventive control problem is generally focused on the optimal power flow constrained by transient stability. In some works of research like [97,98], where DL methods are applied, the constraint conditions of converter control methods are not fully considered. In [99], the current limit value of each VSG is added as the input feature for a DFFNN. By minimizing the total cost of active output adjustment in system, the fast prediction of transient stability is achieved.

In [100], Q-learning is applied to adjust the virtual inertia of each VSG in a 4-VSG system. Two metrics are used to evaluate the performance, one of which is the critical clearing time (CCT) and the other is the percentage of conditions that becomes stable after the parameters are tuned. Compared with traditional methods, the Q-learning method performs better on both of the two metrics. In [101], the authors present an optimal preventive control (OPC) method by tuning the parameters of VSGs, and an RL method with the fault energy reward function is introduced to solve the OPC problem. The simulation results on various systems show that the proposed method can improve the CCT when a fault occurs. In [102], the learning vector quantization (LVQ) network is

applied to distinguish the fault type, which can be performed within 5 ms, and adjust the parameters adaptively according to the type of the fault. The result is compared with some other AI algorithm to demonstrate the effectiveness.

### 5.3. Power Prediction and Scheduling

Power systems with distributed renewable energy provide power for local loads and achieve regional autonomy. With the volatility of renewable energy and the periodic characteristics of industrial and domestic electricity consumption, it is of practical significance that we arrange the power generation plan reasonably. Due to the different generation costs and dispatch costs of different power generation equipment, accurate prediction and reasonable scheduling can bring huge economic benefits to the power system.

In [103], all VSGs output the same frequency when the virtual inertia, the droop coefficient, and the equivalent disturbance are proportional. By adjusting the parameters of each VSG in real time, the oscillation can be suppressed. According to this theory, the author raises an SAC-based model-free method to modify the inertia and droop coefficient, which shows a good robustness under communication interruption and time delay.

Large disturbances and faults can lead to the loss of synchronization (LOS) for VSG control. The influence of virtual inertia varies under different fault conditions and needs different adjustments or modifications within the VSG control framework. Therefore, fault identification and diagnosis are essential to the transient stability of the power system.

The references mentioned in this section and the methods used are shown in Table 2.

**Table 2.** DL and RL in multi-VSG system.

References	Learning Algorithm		Oscillation Suppression	Objective Stability Prediction and Improvement	Power Scheduling	Method	
	DL	RL				Adjust Parameters of each VSG	Predict with Real-Time Data
[84]		✓	✓			✓	
[85]		✓		✓		✓	
[86]	✓		✓			✓	
[86]		✓	✓			✓	
[87]		✓	✓			✓	
[88]		✓	✓			✓	
[89]		✓	✓			✓	
[90]	✓			✓			✓
[91]	✓			✓			✓
[92]	✓			✓			✓
[93]	✓			✓			✓
[94]	✓			✓			✓
[95]	✓			✓			✓
[99]	✓			✓			✓
[100]		✓		✓		✓	
[101]		✓		✓		✓	
[103]		✓			✓	✓	
[102]	✓			✓		✓	

## 6. Discussion and Future Research Trends

Compared with the traditional centralized control strategy, the VSG can serve as a grid-forming control strategy. Compared with droop control, the VSG can provide inertia to the power system under conditions of the high penetration of DGs and power electronic devices. These advantages are destined to keep the VSG-based control strategy the optimal choice in energy internet. And researchers are finding out more solutions with which to overcome the disadvantages, such as difficulties in accurate modeling and parameter tuning, and complex operating conditions, with the help of DL and RL algorithms. It can be drawn from the research mentioned above that DL and RL methods show a powerful ability in improving the performance of the VSG, and, since the research on the implementation of DL and RL in the VSG is still in its infancy, there will be great potential and a lot of room for exploration in this field. Moreover, difficulties and limitations will also arise and need

to be overcome. According to the reviewed literature, some discussions and future research trends are listed as follows:

- **Flexibility.** As mentioned at the beginning of this paper, flexibility is one of the important characteristics of modern power systems. The distributed power supply and loads should be able to be connected or cut off without affecting the conditions of other devices. And, with the further development of distributed power generation, new DGs should be able to connect the system freely, and the parameters of other VSGs should be adjusted automatically if possible, which ensures transient stability under system topology changes.
- **Deployment in real devices or systems.** With many AI-based control methods being proposed, most of them are verified only in a simulation platform or in lab conditions. Few of these methods are applied in real systems. This is due to the high reliability requirements of the power system and incomplete interpretability of DL and RL. However, the combination of an advanced control strategy, VSG, and advanced AI methods, DL and RL, should not remain only at the theoretical level.
- **Efficiency and speed.** Since the VSG control can be applied parallelly in a large system, in which the working conditions are complicated, a powerful DL network or RL agent may need plenty of data with which to train. According to the reviewed research, the training process often takes a relatively long time. A faster training process and a smaller training dataset will bring benefits to the practical deployment of the VSG.
- **Generalization ability.** Since all the DL and RL methods require training processes and all the working conditions of a huge system may not be easy to obtain, a network or agent may perform well on the training datasets but have bad behaviors when a new disturbance occurs. Moreover, topology changes may also lead to failures if operating in a new system with previous parameters. Therefore, the generalization ability of the VSG with DL and RL methods is under great demand. With the strong generalization ability of, the amount of computation and training will be greatly reduced when new systems are deployed.
- **Voltage and reactive power control.** As concluded in Section 3, most of the AI methods for VSG parameter tuning aim to have a better performance in active power and frequency response. The frequency in a system is a global quantity, so the output frequency of each VSG becomes consistent automatically. The active power output is controlled together with the frequency in the same loop, which makes it easier to control. However, due to the existence of line impedance, the deviation in voltage will cause a current circulation between VSGs with different voltage outputs. Since the voltage instability is more complex than the frequency, it is of great importance that we explore the potential of AI methods in voltage control.

**Author Contributions:** Conceptualization, X.D. and J.C.; methodology, X.D.; validation, X.D.; formal analysis, X.D. and J.C.; investigation, X.D.; resources, J.C.; writing—original draft preparation, X.D.; writing—review and editing, X.D. and J.C.; visualization, X.D.; supervision, J.C.; project administration, J.C.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Not applicable.

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

## Abbreviations

A2C	Advantage Actor–Critic
A3C	Asynchronous Advantage Actor–Critic
AI	Artificial Intelligence
BESS	Battery Energy Storage System

CNN	Convolutional Neural Network
DDPG	Deep Deterministic Policy Gradient
DDQN	Double Deep Q-Network
DFFNN	Deep Feed-Forward Neural Network
DG	Distributed Generator
DL	Deep Learning
DPG	Deep Policy Gradient
DNN	Deep Neural Network
DQN	Deep Q-Network
DRL	Deep Reinforcement Learning
EI	Energy Internet
ESS	Energy Storage System
IBDG	Inverter-Based Distributed Generator
IEPE	Institute of Electrical Power Engineering
ILSVRC	ImageNet Large Scale Visual Recognition Competition
KHIs	Kawasaki Heavy Industries
KLD	Kullback–Leibler Divergence
LOS	Loss of Synchronization
LSTM	Long Short-Term Memory
LVQ	Learning Vector Quantization
MDP	Markov Decision Process
ML	Machine Learning
MLP	Multi-Layer Perceptron
MMC	Modular Multilevel Converter
MDHDP	Model-Dependent Heuristic Dynamic Programming
MPGNN	Message Passing Graph Neural Network
OMS	Online Monitoring System
OPC	Optimal Preventive Control
RBM	Restricted Boltzmann Machine
RES	Renewable Energy Source
RL	Reinforcement Learning
RNN	Recurrent Neural Networks
RoCoF	Rate of Change of Frequency
SAC	Soft Actor–Critic
SG	Synchronous Generator
TD	Temporal Difference
TD3	Twin Delayed Deep Deterministic Policy Gradient
TRPO	Trust Region Policy Optimization
VISMA	Virtual Synchronous Machine
VSG	Virtual Synchronous Generator
VSM	Virtual Synchronous Motor

## References

1. Rifkin, J. *The Third Industrial Revolution: How Lateral Power Is Transforming Energy, the Economy, and the World*; Macmillan: London, UK, 2011.
2. Chen, Z.; Liu, Q.; Li, Y.; Liu, S. Discussion on energy internet and its key technology. *J. Power Energy Eng.* **2017**, *5*, 1–9. [[CrossRef](#)]
3. Seneviratne, C.; Ozansoy, C. Frequency response due to a large generator loss with the increasing penetration of wind/PV generation—A literature review. *Renew. Sustain. Energy Rev.* **2016**, *57*, 659–668. [[CrossRef](#)]
4. Xue, Y.; Manjrekar, M.; Lin, C.; Tamayo, M.; Jiang, J.N. Voltage stability and sensitivity analysis of grid-connected photovoltaic systems. In Proceedings of the 2011 IEEE Power and Energy Society General Meeting, Detroit, MI, USA, 24–29 July 2011; IEEE: New York, NY, USA, 2011; pp. 1–7.
5. Crăciun, B.-I.; Kerekes, T.; Séra, D.; Teodorescu, R. Frequency support functions in large PV power plants with active power reserves. *IEEE J. Emerg. Sel. Top. Power Electron.* **2014**, *2*, 849–858. [[CrossRef](#)]
6. Oroojlooy, A.; Hajinezhad, D. A review of cooperative multi-agent deep reinforcement learning. *Appl. Intell.* **2022**, *53*, 13677–13722. [[CrossRef](#)]
7. Beck, H.-P.; Hesse, R. Virtual synchronous machine. In Proceedings of the 2007 9th International Conference on Electrical Power Quality and Utilisation, Barcelona, Spain, 9–11 October 2007; IEEE: New York, NY, USA, 2007; pp. 1–6.



8. Chen, Y.; Hesse, R.; Turschner, D.; Beck, H.-P. Comparison of methods for implementing virtual synchronous machine on inverters. In Proceedings of the International Conference on Renewable Energies and Power Quality, Santiago de Compostela, Spain, 28–30 March 2012.
9. Zhong, Q.-C.; Weiss, G. Synchronverters: Inverters that mimic synchronous generators. *IEEE Trans. Ind. Electron.* **2010**, *58*, 1259–1267. [[CrossRef](#)]
10. Visscher, K.; De Haan, S.W.H. Virtual synchronous machines (VSG's) for frequency stabilisation in future grids with a significant share of decentralized generation. In Proceedings of the CIRED Seminar 2008: SmartGrids for Distribution, Frankfurt, Germany, 23–24 June 2008; IET: Stevenage, UK, 2008.
11. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
12. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
13. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [[CrossRef](#)]
14. Zeiler, M.D.; Fergus, R. Visualizing and understanding convolutional networks. In Proceedings of the Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Proceedings, Part I 13. Springer: Berlin/Heidelberg, Germany, 2017.
15. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
16. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015.
17. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
18. Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014.
19. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.
20. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 91–99. [[CrossRef](#)]
21. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
22. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [[CrossRef](#)]
23. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [[CrossRef](#)] [[PubMed](#)]
24. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)] [[PubMed](#)]
25. Chung, J.; Gulcehre, C.; Cho, K.; Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv* **2014**, arXiv:1412.3555.
26. Hihi, S.; Bengio, Y. Hierarchical recurrent neural networks for long-term dependencies. *Adv. Neural Inf. Process. Syst.* **1995**, *8*, 493–499.
27. Schuster, M.; Paliwal, K.K. Bidirectional recurrent neural networks. *IEEE Trans. Signal Process.* **1997**, *45*, 2673–2681. [[CrossRef](#)]
28. Graves, A.; Fernández, S.; Schmidhuber, J. Multi-dimensional recurrent neural networks. In Proceedings of the International Conference on Artificial Neural Networks, Porto, Portugal, 9–13 September 2007; Springer: Berlin/Heidelberg, Germany, 2007.
29. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6000–6010.
30. Yin, W.; Schütze, H.; Xiang, B.; Zhou, B. Abcnn: Attention-based convolutional neural network for modeling sentence pairs. *Trans. Assoc. Comput. Linguist.* **2016**, *4*, 259–272. [[CrossRef](#)]
31. Liu, B.; Lane, I. Attention-based recurrent neural network models for joint intent detection and slot filling. *arXiv* **2016**, arXiv:1609.01454.
32. Yang, M.; Tu, W.; Wang, J.; Xu, F.; Chen, X. Attention based LSTM for target dependent sentiment classification. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
33. Bellman, R. Dynamic programming. *Science* **1966**, *153*, 34–37. [[CrossRef](#)]
34. Howard, R.A. *Dynamic Programming and Markov Processes*; MITP: Cambridge, MA, USA, 1960.
35. Watkins, C.J.C.H. *Learning from Delayed Rewards*; King's College: Cambridge, UK, 1989.
36. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; MIT Press: Cambridge, MA, USA, 2018.
37. Li, C. Applications of Reinforcement Learning in Three-phase Grid-connected Inverter. In Proceedings of the 2023 IEEE 13th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Qinhuangdao, China, 11–14 July 2023; pp. 580–585.
38. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M. Playing atari with deep reinforcement learning. *arXiv* **2013**, arXiv:1312.5602.
39. Wang, X.; Wang, S.; Liang, X.; Zhao, D.; Huang, J.; Xu, X.; Dai, B.; Miao, Q. Deep Reinforcement Learning: A Survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *35*, 5064–5078. [[CrossRef](#)] [[PubMed](#)]

40. Van Hasselt, H.; Guez, A.; Silver, D. Deep reinforcement learning with double q-learning. In Proceedings of the AAAI Conference on Artificial Intelligence, Phoenix, AZ, USA, 12–17 February 2016.
41. Wang, Z.; Schaul, T.; Hessel, M.; Hasselt, H.; Lanctot, M.; Freitas, N. Dueling network architectures for deep reinforcement learning. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; PMLR: New York, NY, USA, 2016.
42. Schaul, T.; Quan, J.; Antonoglou, I.; Silver, D. Prioritized experience replay. *arXiv* **2015**, arXiv:1511.05952.
43. Schulman, J.; Levine, S.; Abbeel, P.; Jordan, M.; Moritz, P. Trust region policy optimization. In Proceedings of the International Conference on Machine Learning, Lille, France, 6–11 July 2015; PMLR: New York, NY, USA, 2015.
44. Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; Klimov, O. Proximal policy optimization algorithms. *arXiv* **2017**, arXiv:1707.06347.
45. Silver, D.; Lever, G.; Heess, N.; Degris, T.; Wierstra, D.; Riedmiller, M. Deterministic policy gradient algorithms. In Proceedings of the International Conference on Machine Learning, Beijing, China, 21–26 June 2014; PMLR: New York, NY, USA, 2014.
46. Lillicrap, T.P.; Hunt, J.J.; Pritzel, A.; Heess, N.; Erez, T.; Tassa, Y.; Silver, D.; Wierstra, D. Continuous control with deep reinforcement learning. *arXiv* **2015**, arXiv:1509.02971.
47. Mnih, V.; Badia, A.P.; Mirza, M.; Graves, A.; Lillicrap, T.; Harley, T.; Silver, D.; Kavukcuoglu, K. Asynchronous methods for deep reinforcement learning. In Proceedings of the International Conference on Machine Learning, New York, NY, USA, 19–24 June 2016; PMLR: New York, NY, USA, 2016.
48. Haarnoja, T.; Zhou, A.; Abbeel, P.; Levine, S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In Proceedings of the International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; PMLR: New York, NY, USA, 2018.
49. Fujimoto, S.; Hoof, H.; Meger, D. Addressing function approximation error in actor-critic methods. In Proceedings of the International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018; PMLR: New York, NY, USA, 2018.
50. Alipoor, J.; Miura, Y.; Ise, T. Power System Stabilization Using Virtual Synchronous Generator With Alternating Moment of Inertia. *IEEE J. Emerg. Sel. Top. Power Electron.* **2015**, *3*, 451–458. [[CrossRef](#)]
51. Li, D.; Zhu, Q.; Lin, S.; Bian, X.Y. A Self-Adaptive Inertia and Damping Combination Control of VSG to Support Frequency Stability. *IEEE Trans. Energy Convers.* **2017**, *32*, 397–398. [[CrossRef](#)]
52. Wang, F.; Zhang, L.; Feng, X.; Guo, H. An Adaptive Control Strategy for Virtual Synchronous Generator. *IEEE Trans. Ind. Appl.* **2018**, *54*, 5124–5133. [[CrossRef](#)]
53. Xu, Q.; Dragicevic, T.; Xie, L.; Blaabjerg, F. Artificial Intelligence-Based Control Design for Reliable Virtual Synchronous Generators. *IEEE Trans. Power Electron.* **2021**, *36*, 9453–9464. [[CrossRef](#)]
54. Zeng, W.; Xiong, J.; Qi, Z. DCNN-based virtual synchronous generator control to improve frequency stability of PV-ESS station. In Proceedings of the 2022 12th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER), Baishan, China, 27–31 July 2022; pp. 861–865.
55. Zhang, K.; Zhang, C.; Xu, Z.; Ye, S.; Liu, Q.; Lu, Z. A Virtual Synchronous Generator Control Strategy with Q-Learning to Damp Low Frequency Oscillation. In Proceedings of the 2020 Asia Energy and Electrical Engineering Symposium (AEEES), Chengdu, China, 28–30 March 2020.
56. Wu, W.; Guo, F.; Ni, Q.; Liu, X.; Qiu, L.; Fang, Y. Deep Q-Network based Adaptive Robustness Parameters for Virtual Synchronous Generator. In Proceedings of the 2022 IEEE Transportation Electrification Conference and Expo, Asia-Pacific (ITEC Asia-Pacific), Hangzhou, China, 28–31 October 2022; pp. 1–4.
57. Li, Y.; Gao, W.; Yan, W.; Huang, S.; Wang, R.; Gevorgian, V.; Gao, D. Data-driven Optimal Control Strategy for Virtual Synchronous Generator via Deep Reinforcement Learning Approach. *J. Mod. Power Syst. Clean Energy* **2021**, *9*, 919–929. [[CrossRef](#)]
58. Qiu, J.; Yang, H.; Zhang, J.; Gao, J.; Jiang, T.; Gao, Q.; Chen, J.; Xu, G. Parameter tuning of new type energy virtual synchronous generator based on physics-informed reinforcement learning. In Proceedings of the 2023 8th International Conference on Power and Renewable Energy (ICPRE), Shanghai, China, 22–25 September 2023; IEEE: New York, NY, USA, 2023.
59. Liu, P.; Bi, Y.; Liu, C. Data-based intelligent frequency control of VSG via adaptive virtual inertia emulation. *IEEE Syst. J.* **2021**, *16*, 3917–3926. [[CrossRef](#)]
60. Skiparev, V.; Belikov, J.; Petlenkov, E. Reinforcement learning based approach for virtual inertia control in microgrids with renewable energy sources. In Proceedings of the 2020 IEEE PES Innovative Smart Grid Technologies Europe (ISGT-Europe), Amsterdam, The Netherlands, 25–28 October 2020; IEEE: New York, NY, USA, 2020.
61. Xiong, K.; Hu, W.; Zhang, G.; Zhang, Z.; Chen, Z. Deep reinforcement learning based parameter self-tuning control strategy for VSG. *Energy Rep.* **2022**, *8*, 219–226. [[CrossRef](#)]
62. Oboreh-Snapps, O.; She, B.; Fahad, S.; Chen, H.; Kimball, J.; Li, F.; Cui, H.; Bo, R. Virtual Synchronous Generator Control Using Twin Delayed Deep Deterministic Policy Gradient Method. *IEEE Trans. Energy Convers.* **2023**, *39*, 214–228. [[CrossRef](#)]
63. Yang, M.; Wu, X.; Loveth, M.C. A Deep Reinforcement Learning Design for Virtual Synchronous Generators Accommodating Modular Multilevel Converters. *Appl. Sci.* **2023**, *13*, 5879. [[CrossRef](#)]
64. Benhmidouch, Z.; Moufid, S.; Ait-Omar, A.; Abbou, A.; Laabassi, H.; Kang, M.; Chatri, C.; Ali, I.H.O.; Bouzekri, H.; Baek, J. A novel reinforcement learning policy optimization based adaptive VSG control technique for improved frequency stabilization in AC microgrids. *Electr. Power Syst. Res.* **2024**, *230*, 110269. [[CrossRef](#)]

65. Liu, C.; Chu, Z.; Duan, Z.; Zhang, Y. VSG-Based Adaptive Optimal Frequency Regulation for AC Microgrids With Nonlinear Dynamics. *IEEE Trans. Autom. Sci. Eng.* **2024**, 1–12. [[CrossRef](#)]
66. Afifi, M.A.; Marei, M.I.; Mohamad, A.M. Reinforcement Learning Approach with Deep Deterministic Policy Gradient DDPG-Controlled Virtual Synchronous Generator for an Islanded Microgrid. In Proceedings of the 2023 24th International Middle East Power System Conference (MEPCON), Mansoura, Egypt, 19–21 December 2023; IEEE: New York, NY, USA, 2023.
67. Shuai, H.; She, B.; Wang, J.; Li, F. Safe Reinforcement Learning for Grid-Forming Inverter Based Frequency Regulation with Stability Guarantee. *J. Mod. Power Syst. Clean Energy* **2024**, 1–8. [[CrossRef](#)]
68. Wang, Z.; Yu, Y.; Gao, W.; Davari, M.; Deng, C. Adaptive, optimal, virtual synchronous generator control of three-phase grid-connected inverters under different grid conditions—An adaptive dynamic programming approach. *IEEE Trans. Ind. Inform.* **2021**, *18*, 7388–7399. [[CrossRef](#)]
69. Saadatmand, S.; Alharkan, H.; Shamsi, P.; Ferdowsi, M. Model Dependent Heuristic Dynamic Programming Approach in Virtual Inertia-Based Grid-Connected Inverters. In Proceedings of the 2020 52nd North American Power Symposium (NAPS), Tempe, AZ, USA, 11–14 April 2021; IEEE: New York, NY, USA.
70. Saadatmand, S.; Shamsi, P.; Ferdowsi, M. Adaptive critic design-based reinforcement learning approach in controlling virtual inertia-based grid-connected inverters. *Int. J. Electr. Power Energy Syst.* **2021**, *127*, 106657. [[CrossRef](#)]
71. Issa, H.; Debusschere, V.; Garbuio, L.; Lalanda, P.; Hadjsaid, N. Artificial Intelligence-Based Controller for Grid-Forming Inverter-Based Generators. In Proceedings of the 2022 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), Novi Sad, Serbia, 10–12 October 2022; pp. 1–6.
72. Wen, Z.; Long, B.; Lin, K.; Wang, S. Equivalent Modeling Based on Long Short-term Memory Neural Network for Virtual Synchronous Generator. In Proceedings of the 2021 33rd Chinese Control and Decision Conference (CCDC), Kunming, China, 22–24 May 2021; pp. 5881–5886.
73. Ghodsi, M.R.; Tavakoli, A.; Samanfar, A.; Sun, Q. Microgrid Stability Improvement Using a Deep Neural Network Controller Based VSG. *Int. Trans. Electr. Energy Syst.* **2022**, *2022*, 7539173. [[CrossRef](#)]
74. Wang, Z.; Wang, Y.; Davari, M.; Blaabjerg, F. An Effective PQ-Decoupling Control Scheme Using Adaptive Dynamic Programming Approach to Reducing Oscillations of Virtual Synchronous Generators for Grid Connection With Different Impedance Types. *IEEE Trans. Ind. Electron.* **2023**, *71*, 3763–3775. [[CrossRef](#)]
75. Groß, D.; Bolognani, S.; Poolla, B.K.; Dörfler, F. Increasing the resilience of low-inertia power systems by virtual inertia and damping. In Proceedings of the IREP'2017 Symposium, Espinho, Portugal, 27 August 27–1 September 2017; International Institute of Research and Education in Power System Dynamics: Bergen, Norway, 2017.
76. Poolla, B.K.; Bolognani, S.; Dorfler, F. Optimal Placement of Virtual Inertia in Power Grids. *IEEE Trans. Autom. Control* **2017**, *62*, 6209–6220. [[CrossRef](#)]
77. Pournazarian, B.; Sangrody, R.; Lehtonen, M.; Gharehpetian, G.B.; Pouresmaeil, E. Simultaneous Optimization of Virtual Synchronous Generators Parameters and Virtual Impedances in Islanded Microgrids. *IEEE Trans. Smart Grid* **2022**, *13*, 4202–4217. [[CrossRef](#)]
78. Sun, P.; Yao, J.; Zhao, Y.; Fang, X.; Cao, J. Stability Assessment and Damping Optimization Control of Multiple Grid-connected Virtual Synchronous Generators. *IEEE Trans. Energy Convers.* **2021**, *36*, 3555–3567. [[CrossRef](#)]
79. Shi, M.; Chen, X.; Zhou, J.; Chen, Y.; Wen, J. Frequency Restoration and Oscillation Damping of Distributed VSGs in Microgrid With Low Bandwidth Communication. *IEEE Trans. Smart Grid* **2020**, *12*, 1011–1021. [[CrossRef](#)]
80. Xu, Y.; Zhang, W.; Chow, M.Y.; Sun, H.; Gooi, H.B.; Peng, J. A Distributed Model-Free Controller for Enhancing Power System Transient Frequency Stability. *IEEE Trans. Ind. Inform.* **2019**, *15*, 1361–1371. [[CrossRef](#)]
81. Chen, M.; Zhou, D.; Blaabjerg, F. Active Power Oscillation Damping Based on Acceleration Control in Paralleled Virtual Synchronous Generators System. *IEEE Trans. Power Electron.* **2021**, *36*, 9501–9510. [[CrossRef](#)]
82. Fu, S.; Sun, Y.; Liu, Z.; Hou, X.; Han, H.; Su, M. Power oscillation suppression in multi-VSG grid with adaptive virtual inertia. *Int. J. Electr. Power Energy Syst.* **2022**, *135*, 107472. [[CrossRef](#)]
83. Liu, J.; Miura, Y.; Bevrani, H.; Ise, T. Enhanced Virtual Synchronous Generator Control for Parallel Inverters in Microgrids. *IEEE Trans. Smart Grid* **2017**, *8*, 2268–2277. [[CrossRef](#)]
84. Yang, Q.; Yan, L.; Chen, X.; Chen, Y.; Wen, J. A Distributed Dynamic Inertia-Droop Control Strategy Based on Multi-Agent Deep Reinforcement Learning for Multiple Paralleled VSGs. *IEEE Trans. Power Syst.* **2023**, *38*, 5598–5612. [[CrossRef](#)]
85. Wang, H.; Zhang, G.; Hu, W.; Cao, D.; Li, J.; Xu, S.; Xu, D.; Chen, Z. Artificial intelligence based approach to improve the frequency control in hybrid power system. *Energy Rep.* **2020**, *6*, 174–181. [[CrossRef](#)]
86. Gupta, P.; Pal, A.; Vittal, V. Coordinated Wide-Area Damping Control Using Deep Neural Networks and Reinforcement Learning. *IEEE Trans. Power Syst.* **2022**, *37*, 365–376. [[CrossRef](#)]
87. Skiparev, V.; Nosrati, K.; Petlenkov, E.; Belikov, J. Reinforcement Learning Based Virtual Inertia Control of Multi-Area Microgrids. Available online: <https://ssrn.com/abstract=4449057> (accessed on 15 May 2023).
88. Oboreh-Snapps, O.; Strathman, S.A.; Saelens, J.; Fernandes, A.; Kimball, J.W. Addressing Reactive Power Sharing in Parallel Inverter Islanded Microgrid Through Deep Reinforcement Learning. In Proceedings of the 2024 IEEE Applied Power Electronics Conference and Exposition (APEC), Long Beach, CA, USA, 25–29 February 2024; IEEE: New York, NY, USA, 2024.
89. Jiang, J.; Lu, Y.; Pan, T.; Yang, W.; Wang, Y.; Dou, Z. Research on Frequency Optimization Strategy of Multi Virtual Synchronous Generator Microgrid Based on Q-Reinforcement Learning. *Int. J. Focus. Innov. Commun. Intell. Control* **2023**, *19*, 1179.

90. Sun, K.; Likhate, S.; Vittal, V.; Kolluri, V.S.; Mandal, S. An Online Dynamic Security Assessment Scheme Using Phasor Measurements and Decision Trees. *IEEE Trans. Power Syst.* **2007**, *22*, 1935–1943. [[CrossRef](#)]
91. Gupta, A.; Gurrala, G.; Sastry, P.S. An Online Power System Stability Monitoring System Using Convolutional Neural Networks. *IEEE Trans. Power Syst.* **2019**, *34*, 864–872. [[CrossRef](#)]
92. Zhao, H.; Shuai, Z.; Shen, Y.; Chen, H.; Zhao, F.; Shen, X. Online Transient Stability Assessment Method for Microgrid with Multiple Virtual Synchronous Generators Based on Deep Learning. *Autom. Electr. Power Syst.* **2022**, *46*, 109–117.
93. Zhen, J.; Huang, S.; Jia, X.; Du, H. Transient Stability Evaluation of Independent Microgrid based on Message Passing Graph Neural Network. In Proceedings of the 2023 6th International Conference on Electronics Technology (ICET), Chengdu, China, 12–15 May 2023; pp. 861–866.
94. Shen, Y.; Shuai, Z.; Shen, C.; Shen, X.; Ge, J. Transient Angle Stability Prediction of Virtual Synchronous Generator Using LSTM Neural Network. In Proceedings of the 2021 IEEE Energy Conversion Congress and Exposition (ECCE), Vancouver, BC, Canada, 10–14 October 2021; pp. 3383–3387.
95. Zhang, R.; Shuai, Z.; Shen, Y. Data-Mechanism Joint Driven Short-term Voltage Stability Assessment for Islanded Microgrid. In Proceedings of the 2023 26th International Conference on Electrical Machines and Systems (ICEMS), Zhuhai, China, 5–8 November 2023; pp. 4147–4152.
96. Gan, D.Q.; Xin, H.H.; Wang, J.; Du, Z.; Li, Y. Progress in Transient Stability Preventive Control and Optimization. *Autom. Electr. Power Syst.* **2004**, *28*, 1–7.
97. Su, T.; Liu, Y.; Shen, X.; Liu, T.; Qiu, G.; Liu, J. Deep Learning-driven Evolutionary Algorithm for Preventive Control of Power System Transient Stability. *Proc. CSEE* **2020**, *40*, 3813–3824.
98. Yang, Y.; Liu, Y.B.; Liu, J.Y.; Huang, Z.; Liu, T.J.; Qiu, G. Preventive Transient Stability Control Based on Neural Network Security Predictor. *Power Syst. Technol.* **2018**, *42*, 4076–4084.
99. Lin, Y.; Ni, S.; Zhen, J.; Gao, Z. Preventive Control Method for Transient Stability of Independent Microgrid Based on Deep Learning. In Proceedings of the 2022 5th International Conference on Energy, Electrical and Power Engineering (CEEPE), Chongqing, China, 22–24 April 2022; pp. 705–711.
100. Liu, Z.; Zhang, Z. Reinforcement learning-based parameter tuning for virtual synchronous machine on grid transient stability enhancement. In Proceedings of the IECON 2020 the 46th Annual Conference of the IEEE Industrial Electronics Society, Singapore, 18–21 October 2020; IEEE: New York, NY, USA, 2020.
101. Huang, X.; Gwak, J.-Y.; Yu, L.; Zhang, Z.; Cui, H. Transient Stability Preventive Control via Tuning the Parameters of Virtual Synchronous Generators. In Proceedings of the 2023 IEEE Power & Energy Society General Meeting (PESGM), Orlando, FL, USA, 16–20 July 2023; IEEE: New York, NY, USA, 2023.
102. Han, H.; Luo, S.; Chen, S.; Yuan, L.; Shi, G.; Yang, Y.; Fei, L. A transient stability enhancement framework based on rapid fault-type identification for virtual synchronous generators. *Int. J. Electr. Power Energy Syst.* **2024**, *155*, 109545. [[CrossRef](#)]
103. Lin, Y.; Ni, S.; Wang, L. Optimal Scheduling of Virtual Synchronous Independent Microgrid Considering Frequency Stability Constraints. In Proceedings of the 2022 5th International Conference on Energy, Electrical and Power Engineering (CEEPE), Chongqing, China, 22–24 April 2022; pp. 788–794.

**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.