Turkish Journal of Agriculture and Forestry

Volume 47 | Number 5 Article 13

10-16-2023

Optimization of VAE-CGAN structure for missing time-series data complementation of UAV jujube garden aerial surveys

SHUNKANG LING

NIANYI WANG

JINGBIN LI

LONGPENG DING

Follow this and additional works at: https://journals.tubitak.gov.tr/agriculture

 Part of the Agriculture Commons, and the Forest Sciences Commons

Recommended Citation

LING, SHUNKANG; WANG, NIANYI; LI, JINGBIN; and DING, LONGPENG (2023) "Optimization of VAE-CGAN structure for missing time-series data complementation of UAV jujube garden aerial surveys," *Turkish Journal of Agriculture and Forestry*: Vol. 47: No. 5, Article 13. https://doi.org/10.55730/1300-011X.3124
Available at: https://journals.tubitak.gov.tr/agriculture/vol47/iss5/13

This Article is brought to you for free and open access by TÜBİTAK Academic Journals. It has been accepted for inclusion in Turkish Journal of Agriculture and Forestry by an authorized editor of TÜBİTAK Academic Journals. For more information, please contact academic.publications@tubitak.gov.tr.

Turkish Journal of Agriculture and Forestry

http://journals.tubitak.gov.tr/agriculture/

Research Article

Turk J Agric For
(2023) 47: 746-760
© TÜBİTAK
doi:10.55730/1300-011X.3124

Optimization of VAE-CGAN structure for missing time-series data complementation of UAV jujube garden aerial surveys

Shunkang LING[1], Nianyi WANG[1], Jingbin LI[1,2,*], Longpeng DING[1,2]

[1]College of Mechanical and Electrical Engineering, Shihezi University, Shihezi, China

[2]Key Laboratory of Modern Agricultural Machinery of Xinjiang Production and Construction Corps, Shihezi, China

Received: 06.06.2023 • **Accepted/Published Online:** 08.08.2023 • **Final Version:** 16.10.2023

Abstract: UAV aerial survey technology has been widely used in agricultural production, in the aerial survey mission sensor by signal interference, environmental changes and other factors will have the problem of missing flight data. In order to accurately complement the time-series data, the paper proposes a complementary model based on VAE-CGAN optimization, which uses a combination of conditional generative adversarial network (CGAN) and Variational Autoencoder (VAE), incorporates QRNN as a regressor for VAE-CGAN reduction, adds ProbSparse Self-attention (PSA) to reduce computational complexity, and uses a new discriminator structure. Comparative experiments on real aerial survey project datasets show that the model is universally applicable on data with different parameter time series missing and outperforms other comparative models in terms of sample generation capability and prediction results with different missing rates.

Key words: Jujube garden aerial survey, time-series data, missing value complementation, variational autoencoder, conditional generation adversarial network

1. Introduction

In recent years, with the rapid development of information and communication technology and driven by the industrial revolution, agriculture is transforming into precision agriculture and smart agriculture (Nie et al., 2022b, 2022c). At the same time, UAVs, as a control platform that combines many advantages such as data collection and imaging, provide many modern production methods and approaches to the agricultural field (Jiang et al., 2022), playing a role in improving accuracy, efficiency and reducing costs (Pereira et al., 2023). Therefore, UAV technology has a wide range of application prospects in the field of agriculture and can further promote the development of agricultural modernization (Rossi et al., 1994).

The introduction of UAV aerial survey technology in agriculture at this stage is a key component of precision agriculture (PA) (Mulla, 2013), which is usually applied to crop growth monitoring, health monitoring, disease detection, yield estimation, and weed management (Radoglou-Grammatikis et al., 2020). However, there are certain limitations of UAVs in the agricultural production and application process, especially their inability to obtain the required flight data in its entirety, which in turn leads to the challenge of missing data (Tsouros et al., 2019):

- Hardware failure: The aerial photography system of the UAV includes cameras, sensors, and other hardware equipment, and its failure or malfunction after longer-term operation by extreme temperature, static electricity, and vibration may result in incomplete or unusable data being collected.
- Terrain and other signal interference: crop growth terrain complex impedes the signal transmission of the UAV flight, resulting in its reception of the signal is limited or interrupted, and therefore may lead to missing data.
- Flight environment restrictions: Most aerial surveys in the agricultural field are flown in low altitude areas, which can be subject to magnetic interference from the environment, and the control of the UAV can be disturbed and the collected data can be affected. Also, bad weather conditions may delay or interrupt the aerial survey work of the UAV.

UAV flight data is an important basis for judging the operational status and performance of the UAV. The absence of flight data may affect the execution of the mission, or cause the flight control system to fail to work properly and lead to misflight or crash. Completing missing data based on flight data is an important means of UAV safety,

* Correspondence: lijingbin@shzu.edu.cn

This work is licensed under a Creative Commons Attribution 4.0 International License.

stability and success rate of performing aerial survey tasks (Kim et al., 2019). Since the subsequent work requires 3D reconstruction of the collected targets using remote sensing technology, in which the collected remote sensing images need to be fused with the corresponding flight data, it is very important to complete the missing flight data for us to achieve a complete collection of agricultural production information. We would like to choose a deep learning method on complex distribution in this part of the work, which can make accurate complements to the missing temporal data (Sarafanov et al., 2020; Zhang et al., 2021).

For the current data missing completion problem, there are some widely used completion algorithms, their workflow and characteristics are as follows: K-NearestNeighbor (KNN) (Fu et al., 2019; Li and Ercisli, 2023) is a data completion method based on distance similarity, its basic idea is to use similarity to find the nearest neighbors of the missing values, and then predict the missing values based on the data of the nearest neighbors. It is simple and easy to use, can cope with high dimensional and large data sets, but for sparser data, it is easy to have local similarity, which leads to inaccurate prediction of missing values. Generative Adversarial Imputation Network (GAIN) (Yoon et al., 2018) is a GAN-based data complementation method, which works by using a generator to learn the potential feature distribution from the observed data, and introducing a hint matrix in the discriminator to let the discriminator better guide the generator to learn the real data distribution. It relies only on the observed data to perform interpolation operations, which can better improve the accuracy of the model, but has disadvantages such as convergence difficulties and training instability. Imputation using Chained Equations (MICE) (Hayati Rezvan et al., 2015) is a regression model-based data complementation method, which works by filling in missing values with constants such as mean and median, making regression predictions for each variable containing missing values, and using the prediction model for the variables to obtain multiple sets of predicted values for multiple simulations and averaging them to obtain the complemented values. Its preprocessing is simple and can handle multiple missing values, while it is easy to configure and widely used, but it is highly dependent on the sample quality, and the noise present in the data set will greatly affect the results when there are too many missing values and more parameters fitted in the data set.

Among deep learning, generative adversarial network (GAN), as a generative model, is one of the most promising methods for complex distributions in recent years (Bousmalis et al., 2017). It uses generators and discriminators for game learning, i.e. it uses two neural networks, and one neural network is trained against the other in an adversarial process to obtain good results. In response to the sample incompleteness problem caused by the above limitations of UAV aerial survey (Yang et al., 2022), GAN is able to generate virtual samples from the direction of expanding the sample data, and expand the number of samples by feature extraction from the real samples. Back in 2019, PengXu et al. (2019) proposed a neural network-based GAN data augmentation method to synthesize fault data to solve the data imbalance problem in fault prediction of pipeline leaks in petrochemical systems. However, the basic GAN often suffers from training instability, gradient disappearance, pattern collapse, etc. So to solve these problems, derivative models of GAN have been proposed one after another, such as Conditional GAN (CGAN) (Pang and Liu, 2020) with certain constraints entered during sample generation, Generative Adversarial Imputation Network (GAIN) (Yoon et al., 2018) can automatically fill in the missing data, Deep Convolutional GAN (DCGAN) (Viola et al., 2021) employs a deep convolutional neural network, Wasserstein GAN (WGAN) (Arjovsky et al., 2017) employs wasserstein distance to portray the gap between the generated and real samples, InfoGAN (Chen et al., 2016) introduces information bottlenecks, CycleGAN (Almahairi et al., 2018) applies a cyclic consistency loss, Self-Attention GAN (Zhang et al., 2019) introduces a self-attentive mechanism, etc. In agricultural applications, regarding magnetized water fertilizer irrigation, Jing Nie et al. (2022a) combined PGD attack with CGAN to capture the distribution of real data more accurately with limited data samples and generate data for predicting liquid magnetization sequence data. Compared to the generator and discriminator of the basic GAN, where the inputs are only random samples and samples, respectively, the inputs of CGAN can have conditions attached to them. The input to the CGAN generator can be a random sample with some additional features of the sample, so that the desired generated sample can be generated more accurately. The discriminator is the input sample and the corresponding features, which are combined to determine the "truthfulness" of the sample.

In response to the fact that the temporal data generated by GAN is random and cannot capture the mapping of generated samples to temporal information, CGAN largely overcomes this drawback. However, even though CGAN introduces conditional information on the input samples, the input of the generator still has the problem of high randomness. To further improve this problem, this paper uses Variational Autoencoder (VAE) to integrate with CGAN, and the advantages of both complement each other. Both VAE and CGAN essentially perform optimization between distributions, while VAE tends to overwrite the original missing sample distribution by the

full sample distribution of the generated model (Osada et al., 2017). The generator of VAE is able to explore the approximate spatial distribution of the input sample data relatively efficiently, and its training efficiency is therefore higher than that of CGAN. So why do not we just use VAE? The shortcoming of VAE is that the samples generated by decoding are often "fuzzy" (Duan, 2023), and it happens that the discriminator of CGAN is needed to discriminate the authenticity of the samples and help VAE to improve the quality of the samples.

In this paper, the VAE-CGAN optimization model is proposed for the missing flight data problem of jujube garden aerial survey, using the encoder of VAE to replace the noise of traditional CGAN with initialization as well as correction; its decoder when reducing data can be seen as the process of generating samples, so it can be used as a generator of CGAN; the discriminator of CGAN is able to solve the problem of ambiguous samples generated by the decoder of VAE while training its own discriminative ability. In addition, the model incorporates quantile regression neural network (QRNN) as a regressor to complete the regression task in this study, and it can update the data period information by adjusting the confidence interval to reduce the burden for VAE-CGAN. The structure of VAE is described in detail in the paper, and a new structure is tried for the discriminator. In order to verify the superiority of the proposed model, the complementary model, which is widely used nowadays, is selected as a comparison in this paper, and the quality of the samples generated by the generator is evaluated by calculating the MSE and RMSE between the generated data and the real data.

The contributions of this paper are listed as follows:

- Combining CGAN and VAE, and introducing regressors in the overall model architecture, a complementary model for missing UAV flight data is proposed.
- A new discriminator structure was used to compare with the normative one and analyze the advantages and disadvantages of both.
- The proposed model's ability to generate samples and the model prediction results are discussed through experimental analysis of the widely used model.

2. Related work

2.1. Conditional generative adversarial network

The basic generative adversarial network (GAN) is composed of a generator and a discriminator. It performs the generation task of GAN by inputting a low-dimensional random vector to the generator to generate the desired high-dimensional samples. The generated new samples are input to the discriminator to score their truthfulness, and the obtained scalar values range from 0 to 1. The more truthful the sample is, the closer the score is to 1. This is the discriminative task of the GAN. The generator and discriminator are mutually opposing and constraining relationships. By iteratively adjusting the parameters of the generator and discriminator networks, the generator continuously self-corrects and optimizes itself from the feedback of the discriminator, and finally generates samples that better match the conditional samples, and the discriminatory ability of the discriminator gradually improves. The objective function for GAN optimization is:

$$\min_G \max_D V(D,G) = E_{x\sim Pdata(x)}[\log D(x)] + E_{z\sim Pdata(z)}\left[\log\left(1 - D(G(z))\right)\right] \quad (1)$$

In equation (1), x is the original sample and z is the random noise input to the generator, then $G(z)$ represents the generated sample.

For the basic GAN, only a random vector, such as Gaussian noise, can be input, but the generated samples cannot be controlled. It is expected that the GAN can guide the generation of samples to improve the problem of GAN being too free, which is conditional generation. The structure of Conditional Generative Adversarial Network (CGAN) is shown in Figure 1. Not only the conditional variable y is introduced in the generator, because it is obviously not enough for the discriminator to score the input samples based on whether they are true or not, this will cause the generator to gradually ignore the conditions during the learning process, so the conditional variable y is also introduced in the discriminator, and finally it uses "whether sample x is true" and "whether the conditional variable y and sample x match" as the score evaluation criteria. Then the objective function of CGAN optimization becomes:

$$\min_G \max_D V(D,G) = E_{x\sim Pdata(x)}[\log D(x|y)] + E_{z\sim Pdata(z)}\left[\log\left(1 - D(G(z|y))\right)\right] \quad (2)$$

2.2. ProbSparse self-attention

Prediction of long sequences requires models with high predictive power, generally we think of using RNNs (Shipmon et al., 2017), most studies also choose LSTMs, GRUs or mix them with CNNs as the structure of generators and discriminators. Recent studies have shown that Transformer has the potential to improve predictive power, i.e. to efficiently capture the exact long-range correlation coupling between output and input. Transformer, as a model based on the Encoder-Decoder framework, is similar to VAE. However, Transformer has some serious problems: higher quadratic time complexity as well as memory usage, and the quadratic computational complexity of the self-attention (SA) mechanism leads to a model with $O(L_Q L_K)$ time complexity. Therefore, in

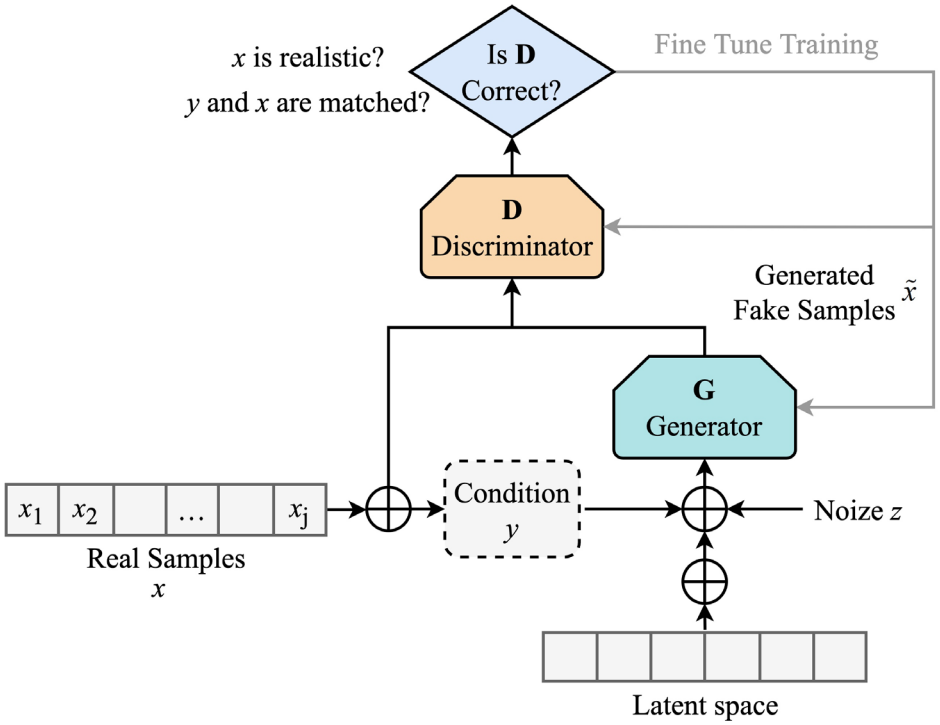

Figure 1. CGAN structure principle.

this paper, ProbSparse Self-attention (PSA) is applied to VAE. PSA controls the computational complexity of SA from $O(l^2)$ to $O(llogl)$ by sampling the difference between the mean distributions of the $U = L_K lnL_Q$ dot product pair calculations, with l being the length of the processed time series, allowing it to accept longer inputs, and its computational complexity is greatly improved compared to traditional SA. In general, PSA improves the missing value interpolation by considering the global characteristics of missing temporal data while reducing the time complexity and computational storage.

2.3. Data preprocessing

The data also needs to be preprocessed before the experiments are formally conducted, with the raw flight data information exported from the ground station side. As time series data segments, data segments with equal time intervals need to be intercepted according to the characteristics of time-series data. Influenced by the limitation of UAV battery life, regional wind gusts, jujube garden aerial survey tasks and safe return, the statistics of flight data found that 100% of jujube garden aerial survey tasks were less than 25 min long, and 91% of the tasks were concentrated in 15 to 25 min; setting the 95% quantile as the maximum length, the value was calculated to be 19 min, which was selected as the length of the time series data segment and carried out data preprocessing.

The specific operation steps are: firstly, judge the length of all data segments, and do the deletion process for the voyage whose duration exceeds 19 min, i.e. delete the data after the moment of more than 19 min; do the padding process for the sample of voyage whose duration is less than 19 min, i.e. perform the 0-filling operation for the empty segments of data whose duration is less than 19 min, so that their duration reaches 19 min.

The purpose of the above data preprocessing practice is that when selecting the sample length as the interception criterion, the effect of model training needs to be taken into account while ensuring the integrity of the sample. If the selected criterion length is too long, too many zeros will be added when padding the shorter samples, which may lead to the model not being trained correctly or overfitting phenomenon, affecting the generalization ability of the model. On the contrary, if the selected standard length is too short, it will lead to a serious lack of sample information, which will also have a negative impact on the performance of the model (Li et al., 2022). Therefore, when selecting the criterion length, this paper takes into account the data distribution, model architecture and task requirements to determine the most appropriate sample length to fully utilize the data information to verify the model performance.

3. Models architecture and methodology

3.1. Variational autoencoder and generator

The rationale for VAE's ability to generate data similar to the original dataset lies in its desire to train a model $x = G(z)$ that maps the original probability distribution to the probability distribution of the training set, i.e. it aims to perform a transformation between distributions so that it learns the probability distribution of the original dataset well. So VAE introduces the hidden variable z and assumes that $P(z|x)$ is normally distributed. The resampling method was applied in the process of sampling, and the mean value calculation was done on the basis of the hidden variable z, while the variance calculation was done by sampling from the standard normal distribution, which was recovered to the original distribution by the mean and variance. Although the noise (i.e. variance) in the output through the neural network makes reconstruction more difficult, eventually the model is reconstructed as well as possible so that the variance is 0. The randomness of the model is thus gradually reduced, and the encoder degenerates so that only the mean value is obtained after sampling, and the noise loses its effect. So to prevent the noise from being zero and to ensure the generative power of the model, VAE makes $P(z|x)$ align to the standard normal distribution and measures the KL dispersion between the specific normal distribution and the standard normal distribution as an additional loss, which consists of both mean calculation and variance calculation, expressed as Equation (3):

$$L_{\mu,\sigma^2} = \frac{1}{2}\left(\mu_{(i)}^2 + \sigma_{(i)}^2 - \log \sigma_{(i)}^2 - 1\right) \quad (3)$$

Based on this model architecture, adding PSA to the encoder and decoder can capture the features in the timing data, and by training this model, the final generated timing data matrix can effectively solve the missing in the original data, and the internal structure of VAE is shown in Figure 2.

3.1.1. Encoder

The encoder in VAE compresses the original input temporal data matrix into a low-dimensional vector, and it obtains the hidden variable z as follows:

a. Set Batch_Size = m, randomly draw m samples $\{x_1,...,x_m\}$ from the original sample set as the temporal data matrix, and select its corresponding mask matrix as $\{M_1,...,M_m\}$;

b. Input $\{x_1,...,x_m\}$ into the encoder unit of the VAE and into the PSA layer to obtain

the context vector as the representation;

c. The output of the previous step into the global average pooling layer can be directly implemented to reduce the dimensionality and obtain a vector;

d. The output of the previous step enters into two fully connected layers to obtain the mean and variance vectors respectively, and the specific normal distribution is determined by the mean and variance vectors;

e. After specifying the parameters of the normal distribution, the hidden variable z is obtained by sampling from that normal distribution.

3.1.2. Decoder

The decoder in VAE maps the low-dimensional vector z output from the encoder to the high-dimensional samples, and it obtains the complete temporal data as follows:

a. The hidden variable z obtained from the encoder is input to the decoder unit of VAE and the KL dispersion L_{μ,σ^2} between the current specific normal distribution and the standard normal distribution is calculated;

b. The hidden variable z is guided by the data period information y. A reshape operation is performed to reconstruct the array structure to prevent data distortion and transform the multidimensional array form of the tensor into the form of a parametric shape;

c. The output of the previous step goes to the PSA layer to obtain the context vector as a representation;

d. The output of the previous step goes to the fully connected layer to obtain the complete time-series data matrix;

e. With the generated time series data, the mean square error loss LMSE between it and the input original series data is calculated; the total loss $Loss = L_{\mu,\sigma^2} + L_{MSE}$, and the parameters of VAE are updated according to the loss function back propagation.

3.2. Discriminator structures

RNN is often used in processing time series data, and LSTM (Hochreiter and Schmidhuber, 1997) and GRU are used as its variants to solve its gradient vanishing problem to some extent (Wang et al., 2022). In this paper, LSTM is chosen in the structure of the discriminator because it is theoretically superior to GRU, although the network structure of GRU is simpler than it, but in fact the framework structure of the discriminator is not complicated and the simplification degree of part of the network structure is not required. While considering the data for time period dependence, the long-term nonlinear correlation cannot be ignored, so CNN is added to the structure of the discriminator, and the combined model based on CNN-LSTM enhances the discriminative ability of the discriminator.

In the following two discriminator structures, a single-layer one-dimensional CNN with 3 × 3 convolutional kernel size and 64 convolutional kernels is chosen, and the number of neurons in the single-layer LSTM is 128 and 256, respectively. The activation function is LeakyReLU, and the negative gradient parameter α is taken as 0.05. This setting can avoid the problem that the input is less than zero and thus the gradient is zero, which eventually leads to its weights not getting updated (i.e. dead neurons). The input real sample x and the generated sample $\hat{x}$ are

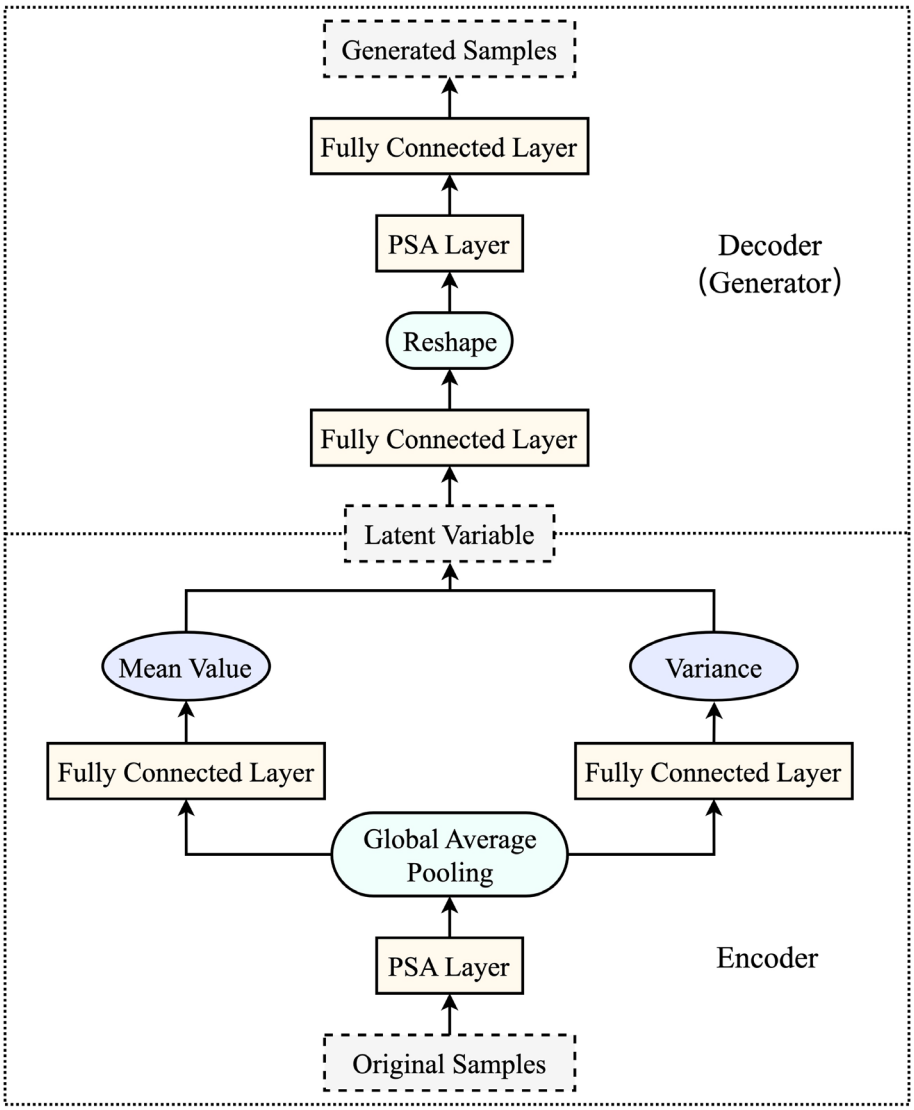

Figure 2. VAE internal structure.

processed through the previous network layers, and finally enter a fully connected layer containing only 1 neuron, where the input vector is subjected to a dimensionality reduction operation and the output dimension is set to 1. A Sigmoid activation function is required to restrict the final output to a value between 0 and 1. The hidden state of the time series data is then compressed into a one-dimensional value by a fully connected layer and mapped to the (0,1) interval by a Sigmoid activation function, which represents the discriminator's "score" for its input samples.

3.2.1. Normal structure

The structure of the discriminator, which is generally more common in CGAN, is shown in Figure 3. The true sample x, the generated sample $\hat{x}$, and the data cycle information y pass through their respective networks, with the true sample x and the generated sample $\hat{x}$ passing through a two-layer CNN-LSTM neural network, and the data cycle information y, which is only temporally relevant, passing through an LSTM neural network, and these two networks process the features through their respective embedding layers, and then the intrinsic properties of the features captured by these two embedding layers are all input to another LSTM neural network to obtain a score to measure "whether sample x is true". The intrinsic properties of the features captured by both embedding layers are then fed into the other LSTM neural network to obtain a score to measure "whether sample x is true" and "whether the conditional variables y and sample x match".

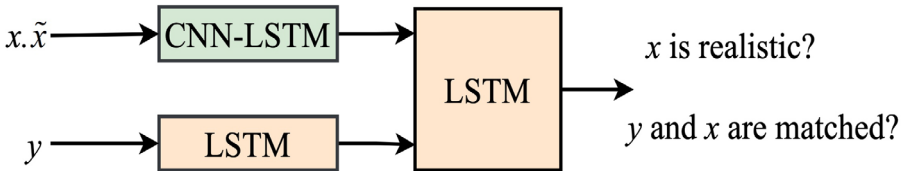

Figure 3. Normal structure.

3.2.2. Proposed structure

In this paper, we propose to use a less common structure of CGAN, as shown in Figure 4. The true sample x, the generated sample $\hat{x}$ are first passed through a CNN-LSTM neural network, which outputs two separate pieces of information, one for measuring "whether sample x is true" and the other for the features obtained from these samples through the embedding layer, and then these features and the data cycle information y are are then fed into the LSTM neural network and a score is output to measure "whether the condition variable y matches the sample x".

This structure is theoretically more reasonable than the common structure because it can clearly know the distribution of the two measures "whether sample x is true" and "whether the condition variable y matches sample x" in the score, whereas the common structure does not clearly give the reason for the score by the final. For example, if the discriminator scores low, it does not know whether the sample is not true enough or the condition variable does not match enough.

In general, the advantage of this CGAN discriminator structure is that it can consider both the authenticity of the generated samples and the degree of matching with the condition variables, rather than just judging the authenticity of the samples. This structure of CNN-LSTM neural network can extract the features of the samples and transform these features into a more semantically informative representation through the embedding layer. This allows the discriminator to have a stronger semantic understanding, thus improving the match between the generated samples and the condition variables. In addition, this structure can also take into account the time-period information in the data and better reflect the characteristics of time-series data. Therefore, compared with the conventional CGAN discriminator structure, this structure can better match the condition variables while ensuring the authenticity of the generated samples, thus improving the performance of the generated model.

3.3. Overall model architecture

The essential difference between VAE and CGAN determines the different roles they assume in the overall model, and the difference lies in the loss function. VAE applies the pointwise loss function, which does not consider the correlation between samples, so a typical feature of VAE is that the pointwise loss function is often detached from the data stream shape surface, and the lack of global similarity causes the generated samples "fuzzy". While CGAN is the distribution matching loss, it discriminator of the reconstruction loss calculates the distance between the probability stream shape between the generated samples and the original samples, which is closer to the popular surface, and the generated samples will be clearer, even with additional conditions, the distribution matching still has some difficulties, such as the generation of pattern collapse. Pointwise loss function of VAE can compensate this problem, and its encoder is used to do initialization or to do correction, while its decoder is to reduce the spatial distribution of the low-dimensional manifold to the original data, which functions similarly to the generator of CGAN. From this perspective, it seems that the input of the generator (i.e. decoder) is no longer random noise. Therefore, a reasonable combination of VAE and CGAN makes the model accurate and efficient.

Although we hope that the decoder output $\hat{x}$ of VAE should be as close as possible to the original x, i.e. the smaller the loss, the better, the decoded samples are often difficult to reach zero loss, and the smaller the loss does not mean that the samples are more realistic, so a discriminator with "conditions" is needed to discriminate the real degree of the samples. For the prediction of time series data, the time continuity is often defaulted, however, in unsupervised learning, if the time dependence of the sample is ignored, it is easy to make a judgment error. For example, the presence of large angle maneuvers in a flight mission leads to abrupt changes in parameters, and if only the temporal continuity of the data is learned, it is likely that the originally correct data will be misjudged as incorrect data. Using the data period information as the condition information of CGAN, in a sense, transforms the unsupervised learning into supervised learning, which not only guides the decoder to establish the sample-time-dependent relationship when generating samples, and better learns the distribution of physical and chemical property parameters in the time interval where it is located, but also reduces the misjudgment rate of the discriminator. This is because the discriminator not only has to discriminate the authenticity of the generated samples $\hat{x}$, but also has to check whether the time series data match the data period information y, effectively avoiding the occurrence of misclassification.

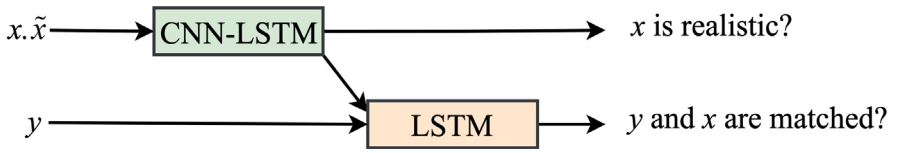

Figure 4. Proposed structure.

By combining CGAN and VAE, this paper embeds a regressor in the overall framework so that the model has the ability to handle regression problems and accomplish the task of completing missing data for the flight task, while being able to guarantee the quality of the generated samples and reduce the burden for the training of discriminators and generators. The regressor uses a quantile regression neural network (QRNN) with three fully-connected layers, which has three major advantages, firstly, in terms of training speed, as QRNN uses convolutional operations and gating structure to reduce the number of repetitive computations and time, so it can be faster and more stable than traditional recurrent neural networks (RNN) or LSTM (Nie et al., 2021). Secondly, the gating structure in QRNN can ignore useless information and select useful information for updating the state vector; finally, in terms of generalization ability, QRNN can learn the intrinsic characteristics of data by using fully connected layers, which can better adapt to different types and sizes of data sets.

Thus, the role of this regressor in the model is to correspond the generated samples $\hat{x}$ to the conditional variables y with a certain confidence level, to drive the generated samples to fall within a certain confidence interval, to relax the requirements of the discriminator and the generator, and to establish a mapping relationship that can be further updated by iteratively updating the conditional variables y and adjusting the confidence interval to improve the quality of the generated samples and the discriminator's ability to discriminate samples. The mapping relationship established can be further updated by iteratively updating the condition variable y and adjusting the confidence interval to improve the quality of samples generated by the generator and the ability of the discriminator to discriminate samples. Finally, the overall model architecture is shown in Figure 5.

4. Experimental methods and results

4.1. Dataset

In order to evaluate the trend prediction performance of the algorithm proposed in this paper, the Mavic3 UAV from DJI was selected as the platform for mounting the panoramic camera, and its vertical and horizontal accuracy was ±0.5 m when the satellite positioning was working normally; the ONE X2 panoramic camera from Insta was mounted under the platform as the jujube tree image acquisition equipment, and its weight was 149 g, as shown in Figure 6. The area of this aerial survey mission was selected from the standardized jujube palm orchards under the jurisdiction and part of the 224th Regiment of the Production and Construction Corps, 14th Division, Kunyu City, Xinjiang Uygur Autonomous Region, China ($37^0$2'~$37^0$21'N, $79^0$15'~$79^0$20'E), as shown in Figure 7. It is located at the northern foothills of Kunlun Mountains, in the southwest of Tarim Basin in Hotan Region, and belongs to the Taklamakan Desert hinterland, where the topography and landscape are dominated by sand dunes are dominant, and the gusts are strong.

In the aerial survey mission of jujube tree image acquisition, there are mission preparation phase, flight mission phase and data postprocessing phase. In the preparation phase, the flight route is designed according to the task area of the jujube garden and the camera parameters are set according to the environment; in the flight mission phase, it includes carrying the UAV device and taking off, flying according to the route that meets the actual sampling requirements, and returning with a safe descent route; in the postprocessing phase, the panoramic images with time-series information collected by the panoramic camera are downloaded, and the data from this flight log in the UAV sensor are downloaded. The two data are used for follow-up work.

In the flight mission phase, to realize the image acquisition of jujube trees, the UAV needs to climb to the center of the jujube tree canopy height to keep hovering, and fix the height for traverse and yaw adjustment to the middle of the jujube tree row, then fly along the jujube tree row in a straight line to the end and traverse to the adjacent jujube tree row, repeat the above actions until the aerial survey mission is completed or the power is low, and finally return safely around the surrounding obstacles. The mission is representative, and the flight route is shown in Figure 8.

The flight data comes from GPS, inertial measurement unit (IMU) and barometer sensor, and the sensor data mainly contains longitude, latitude, altitude, yaw angle, pitch angle, acceleration and other information. In this paper, we mainly choose the position information with the characteristics of time-series trend change as the research object, so we extract four parameters of longitude,

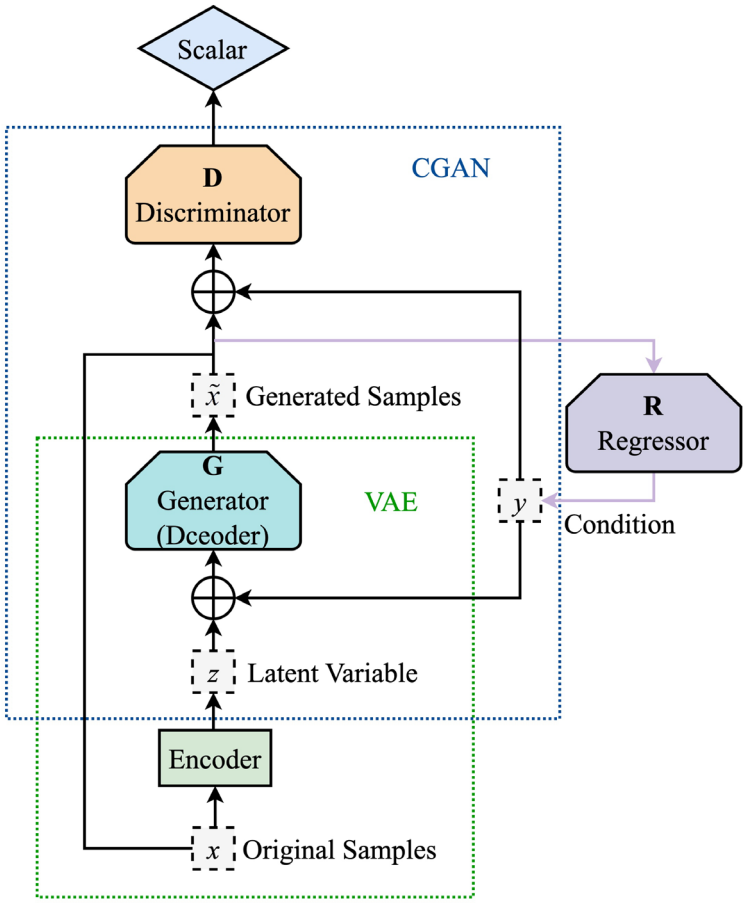

Figure 5. Overall model architecture.

Figure 6. Aerial survey operation drone physical picture.

Figure 7. Standardized orchards in southern Xinjiang.

Figure 8. UAV mission route map.

latitude, altitude information and heading angle as the characteristic data, among which longitude is converted into the east-west deviation from the takeoff point and latitude is converted into the north-south deviation from the takeoff point to reflect the flight trajectory intuitively.

The characteristics of different parameter data sets are influenced by the aerial survey task, the UAV has low acceleration and large distance span in the horizontal direction, so the latitude and longitude data show gentle changes with large amplitude; the UAV keeps orientation between jujube tree rows but needs to turn drastically to change rows at the end, so the heading angle data show intermittent and large sudden changes; because the jujube trees are low in height and the surrounding terrain is flat, the height data show irregular and small changes.

In the experiment, 2400 data segments were selected for testing in this paper, with 600 data segments for each parameter. Before conducting the experiment, we

preprocessed them. Then, the first 450 data segments for each parameter were selected as training samples, the last 50 were selected as validation set, and the last 100 were selected as test set. The total duration of each data segment was preprocessed to a uniform 19 min and made to sample the sensor every 10 s, so that the length of each data segment was 114 points. In this way, we can obtain a large amount of data to analyze and study our experiments.

4.2. Methods

The experimental hardware is an NVIDIA GeForce RTX 2080 with 32 GB of RAM, the software system environment is Windows 10, and the software languages and deep learning frameworks are Python (version 3.7.11) and Pytorch (version 1.9.0).

In this paper, an optimization model with VAE-CGAN as the basic structure is built to address the problem of missing UAV flight data complement, and the model has the following special features:

- The model combines CGAN and VAE, focusing on the VAE decoder as a CGAN generator and combining it with the VAE encoder as a new generator, the main part of its training signal is provided by the VAE reconstruction error target, so its core is a measure of the similarity of the two probability distributions, compared to the original single CGAN generator with a random strategy.
- Model joint training of VAE-CGAN with element-based reconstruction metrics replaced by feature-based discriminators, since sample similarity can be measured by discriminators of CGAN, this paper proposes the use of a new structure of discriminators, which is more theoretically sound.
- The model embeds a regressor in addition to VAE-CGAN as a further optimization, which not only dominates the implementation of the regression prediction task, but also alleviates the training difficulty of VAE-CGAN with its assistance. In addition, the computational complexity is reduced by using ProbSparse Self-attention in the model, and these factors can theoretically improve the accuracy of prediction results as well as the speed of training.

To this end, this paper hopes to explore these characteristics separately through whole experiments on different attribute datasets, and to discuss and analyze the following specific experiments were designed:

- The widely used complementation models are selected for performance comparison with the model designed in this paper, including K-NearestNeighbor (KNN) (Fu et al., 2019; Li and Ercisli, 2023), generic adversarial imputation nets (GAIN), and multiple Imputation using Chained Equations (MICE) (Hayati Rezvan et al., 2015).
- Verify the gap between the generated data and the real value, design suitable performance evaluation indexes, evaluate the prediction results of each model, and analyze the prediction results.

4.3. Generated samples quality

Generally, in order to verify the merits of the generated results, the commonly used performance evaluation metrics are MAE, MAPE, MSE, and RMSE. Due to the existence of the actual value of zero in the dataset, this will lead to the inability to calculate MAPE, and also due to the low sensitivity of MAE to outliers, only the absolute value of the error is considered, and the magnitude of the change of the error is not considered. Therefore, in this paper, Mean Square Error (MSE) and Root Mean Square Error (RMSE) are chosen as the evaluation indicators. MSE is the squared difference between the predicted and true values of the model and calculates their average value, which is more sensitive to large errors, which means that MSE is more sensitive to data with outliers, and it considers 'the smaller the MSE, the higher the accuracy of the model prediction'. The MSE is calculated as shown in Equation (4).

$$MSE = \frac{1}{N}\sum_{i=1}^{n}(\hat{x}_i - x_i)^2 \quad (4)$$

RMSE is the open square of the mean of the square of the average difference between the predicted and true values of the model. The calculation process is similar to MSE, but uses the square root approach to eliminate the magnitude problem. Replacing the units of MSE with the same units as the original target variable makes the error quantification more intuitive and meaningful, and easier to interpret in actual numerical comparisons. The smaller the value of RMSE, the closer the prediction result. The smaller the value of RMSE, the closer the average distance between the prediction result and the true value, and the higher the prediction accuracy of the model. The RMSE is calculated as shown in Equation (5).

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(\hat{x}_i - x_i)^2} \quad (5)$$

where x and $\hat{x}$ represent the original and predicted generated values, respectively, and n is the number of samples.

4.4. Prediction results

In this paper, we compare the generated results of five models, KNN (Chao and Li, 2022), GAIN, MICE, VAE-CGAN, and VAE-CGAN optimization model. Select complete sequence data from the self-constructed dataset and perform complete random deletion processing according to varying degrees of sample missing rates. The reason for this processing is that the type of data missing in the real environment is Missing Completely At Random (MCAR), which means that data missing in the dataset

is completely random and independent of the missing variable itself and other variables. Since the actual self-constructed datasets have a minimum missing rate of 5% and a maximum of 30% of values, where more than 85% of the sample data are missing at less than 20%, the missing rates are set to 5%, 10%, 15%, 20%, and 30%. Data sets for each parameter with different missing rates were data generated by the five models mentioned above, and MAE and RMSE were calculated between them and the original data set, and the final experimental results are shown in Tables 1-4.

The following conclusions can be obtained by analyzing the experimental results:

- The accuracy of each complementation model shows an overall increasing trend with the increase of the missing proportion, which indicates that the missing proportion of data is one of the important factors affecting the accuracy of the complementation, among which the MICE model has a lower accuracy under a large missing rate.
- The performance of the same method varies among data sets with different characteristics, which indicates that there is a relationship between the performance of the model and the changing characteristics of the data set.
- For the problem of missing flight data completion in aerial survey tasks, the VAE-CGAN-based model proposed in this paper outperforms other models in terms of accuracy, generalization and robustness for different types of data sets and missing rates.

5. Conclusion and future work

This paper explores a method that can use incomplete samples for trend prediction based on the generative model of VAE-CGAN. When machine learning is applied to agriculture, the first problem that needs attention is the difficulty of collecting samples, which are affected by the complex environment, whether they are collected as text samples or image samples, there are missing sample data and a small number of complete samples. The method proposed in this paper is studied for jujube garden UAV aerial survey flight data, and the missing temporal data samples are complemented to obtain high quality complete samples, and the regression task of trend prediction is further completed as the optimization link of missing value complementation. The proposed complementation model in this paper provides a reliable solution to the problem of missing UAV flight data, which is important to effectively guarantee the successful completion of UAV aerial survey tasks.

In the application scenario of missing UAV flight data completion, due to the huge flight data and high time

Table 1. MSE, RMSE between longitude generated data and original data.

Miss-ing rate	KNN		GAIN		MICE		VAE-CGAN		VAE-CGAN-QRNN	
	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE
5%	0.512	0.716	0.424	0.651	0.432	0.657	0.412	0.642	0.408	0.639
10%	0.551	0.742	0.451	0.672	0.462	0.680	0.433	0.658	0.43	0.656
15%	0.593	0.770	0.526	0.725	0.528	0.727	0.491	0.701	0.487	0.698
20%	0.631	0.794	0.591	0.769	0.602	0.776	0.557	0.746	0.552	0.743
30%	0.671	0.819	0.628	0.792	0.662	0.814	0.598	0.773	0.594	0.771

Table 2. MSE, RMSE between latitude generated data and original data.

Miss-ing rate	KNN		GAIN		MICE		VAE-CGAN		VAE-CGAN-QRNN	
	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE
5%	0.517	0.719	0.427	0.653	0.435	0.660	0.417	0.646	0.413	0.643
10%	0.556	0.746	0.458	0.677	0.467	0.683	0.438	0.662	0.435	0.660
15%	0.599	0.774	0.529	0.727	0.535	0.731	0.497	0.705	0.494	0.703
20%	0.638	0.799	0.596	0.772	0.607	0.779	0.563	0.750	0.558	0.747
30%	0.679	0.824	0.632	0.795	0.668	0.817	0.61	0.781	0.604	0.777

Table 3. MSE, RMSE between altitude generated data and original data.

Missing rate	KNN		GAIN		MICE		VAE-CGAN		VAE-CGAN-QRNN	
	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE
5%	0.393	0.627	0.398	0.631	0.41	0.640	0.334	0.578	0.332	0.576
10%	0.464	0.681	0.419	0.647	0.436	0.660	0.337	0.581	0.336	0.580
15%	0.532	0.729	0.501	0.708	0.506	0.711	0.471	0.686	0.468	0.684
20%	0.604	0.777	0.559	0.748	0.573	0.757	0.504	0.710	0.498	0.706
30%	0.641	0.801	0.637	0.798	0.649	0.806	0.602	0.776	0.598	0.773

Table 4. MSE, RMSE between heading angle generated data and original data.

Missing rate	KNN		GAIN		MICE		VAE-CGAN		VAE-CGAN-QRNN	
	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE	MSE	RMSE
5%	0.562	0.750	0.443	0.666	0.463	0.680	0.431	0.657	0.429	0.655
10%	0.602	0.776	0.47	0.686	0.496	0.704	0.454	0.674	0.475	0.689
15%	0.645	0.803	0.546	0.739	0.538	0.733	0.513	0.716	0.502	0.709
20%	0.684	0.827	0.612	0.782	0.633	0.796	0.578	0.760	0.574	0.758
30%	0.711	0.843	0.648	0.805	0.699	0.836	0.619	0.787	0.616	0.785

complexity, the model cannot well meet the real-time requirements with limited computational performance, so the design of computationally lightweight missing data completion algorithms becomes one of the important research directions in the future. In addition, because the UAV as a whole, many parameters of flight data are correlated, and if the correlation between different types of data can be combined will lead to a greater improvement in prediction accuracy, however, the existing algorithms usually have problems such as dimensional disaster, excessive volume of algorithms and underfitting when dealing with spatio-temporal high-dimensional data with correlation, which will also become a key problem to be solved in future work.

Acknowledgment
This work was supported by the National Natural Science Foundation of China (52165037), the National Natural Science Foundation of China (52265038) and the Corps Regional Innovation Guidance Program (2021BB003)

References

Almahairi A, Rajeshwar S, Sordoni A, Bachman P, Courville A (2018). Augmented cyclegan: Learning many-to-many mappings from unpaired data. In: International conference on machine learning; pp. 195-204.

Arjovsky M, Chintala S, Bottou L (2017). Wasserstein generative adversarial networks. In: International conference on machine learning; pp. 214-223.

Bousmalis K, Silberman N, Dohan D, Erhan D, Krishnan D (2017). Unsupervised pixel-level domain adaptation with generative adversarial networks. In: Proceedings of the Institute of Electrical and Electronics Engineers conference on computer vision and pattern recognition; pp. 3722-3731.

Chao X, Li Y (2022). Semisupervised Few-Shot Remote Sensing Image Classification Based on KNN Distance Entropy. Institute of Electrical and Electronics Engineers Journal of Selected Topics in Applied Earth Observations and Remote Sensing 15: 8798-8805. https://doi.org/10.1109/JSTARS.2022.3213749

Chen X, Duan Y, Houthooft R, Schulman J, Sutskever I et al. (2016). Infogan: Interpretable representation learning by information maximizing generative adversarial nets. Advances in neural information processing systems 29.

Duan J (2023). Researches Advanced in Image Generation based on Deep Learning. Highlights in Science, Engineering and Technology 39: 413-418. https://doi.org/10.54097/hset.v39i.6561

Fu Y, He HS, Hawbaker TJ, Henne PD, Zhu Z et al. (2019). Evaluating k-Nearest Neighbor (kNN) Imputation Models for Species-Level Aboveground Forest Biomass Mapping in Northeast China. Remote sensing 11 (17): 2005. https://doi.org/10.3390/rs11172005

Hayati Rezvan P, Lee KJ, Simpson JA (2015). The rise of multiple imputation: a review of the reporting and implementation of the method in medical research. BMC medical research methodology 15: 1-14. https://doi.org/10.1186/s12874-015-0022-1

Hochreiter S, Schmidhuber J (1997). Long short-term memory. Neural computation 9 (8): 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Jiang J, Zhang Q, Wang W, Wu Y, Zheng H et al. (2022). MACA: A relative radiometric correction method for multiflight unmanned aerial vehicle images based on concurrent satellite imagery. Institute of Electrical and Electronics Engineers Transactions on Geoscience and Remote Sensing 60: 1-14. https://doi.org/10.1109/TGRS.2022.3158644

Kim T, Ko W, Kim J (2019). Analysis and impact evaluation of missing data imputation in day-ahead PV generation forecasting. Applied Sciences 9 (1): 204. https://doi.org/10.3390/app9010204

Li Y, Ercisli S (2023). Data-efficient crop pest recognition based on KNN distance entropy. Sustainable Computing: Informatics and Systems 38: 100860. https://doi.org/10.1016/j.suscom.2023.100860

Li Y, Gong Y, Zhang Z (2022). Few-shot object detection based on self-knowledge distillation. Institute of Electrical and Electronics Engineers Intelligent Systems. https://doi.org/10.1109/MIS.2022.3205686

Mulla DJ (2013). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. Biosystems engineering 114 (4): 358-371. https://doi.org/10.1016/j.biosystemseng.2012.08.009

Nie J, Wang N, Li J, Wang K, Wang H (2021). Meta-learning prediction of physical and chemical properties of magnetized water and fertilizer based on LSTM. Plant Methods 17 (1): 1-13. https://doi.org/10.1186/s13007-021-00818-2

Nie J, Wang N, Li J, Wang Y, Wang K (2022a). Prediction of liquid magnetization series data in agriculture based on enhanced CGAN. Frontiers in Plant Science: 1883. https://doi.org/10.3389/fpls.2022.929140

Nie J, Wang Y, Li Y, Chao X (2022b). Artificial intelligence and digital twins in sustainable agriculture and forestry: a survey. Turkish Journal of Agriculture and Forestry 46 (5): 642-661. https://doi.org/10.55730/1300-011X.3033

Nie J, Wang Y, Li Y, Chao X (2022c). Sustainable computing in smart agriculture: survey and challenges. Turkish Journal of Agriculture and Forestry 46 (4): 550-566. https://doi.org/10.55730/1300-011X.3025

Osada G, Omote K, Nishide T (2017). Network intrusion detection based on semi-supervised variational auto-encoder. In: Computer Security–ESORICS 2017: 22nd European Symposium on Research in Computer Security, Oslo, Norway, September 11-15, 2017, Proceedings, Part II 22; pp. 344-361.

Pang Y, Liu Y (2020). Conditional generative adversarial networks (CGAN) for aircraft trajectory prediction considering weather effects. In: the American Institute of Aeronautics and Astronautics Scitech 2020 Forum (AIAA Scitech Forum); pp. 1853.

Pereira FRdS, Dos Reis AA, Freitas RG, Oliveira SRdM, Amaral LRd et al. (2023). Imputation of Missing Parts in UAV Orthomosaics Using PlanetScope and Sentinel-2 Data: A Case Study in a Grass-Dominated Area. International Society for Photogrammetry and Remote Sensing International Journal of Geo-Information (ISPRS International Journal of Geo-Information) 12 (2): 41. https://doi.org/10.3390/ijgi12020041

Radoglou-Grammatikis P, Sarigiannidis P, Lagkas T, Moscholios I (2020). A compilation of UAV applications for precision agriculture. Computer Networks 172: 107148. https://doi.org/10.1016/j.comnet.2020.107148

Rossi RE, Dungan JL, Beck LR (1994). Kriging in the shadows: geostatistical interpolation for remote sensing. Remote Sensing of Environment 49 (1): 32-40. https://doi.org/10.1016/0034-4257(94)90057-4

Sarafanov M, Kazakov E, Nikitin NO, Kalyuzhnaya AV (2020). A machine learning approach for remote sensing data gap-filling with open-source implementation: An example regarding land surface temperature, surface albedo and NDVI. Remote sensing 12 (23): 3865. https://doi.org/10.3390/rs12233865

Shipmon DT, Gurevitch JM, Piselli PM, Edwards ST (2017). Time series anomaly detection; detection of anomalous drops with limited features and sparse examples in noisy highly periodic data. arXiv preprint arXiv:1708.03665. https://doi.org/10.48550/arXiv.1708.03665

Tsouros DC, Bibi S, Sarigiannidis PG (2019). A review on UAV-based applications for precision agriculture. Information 10 (11): 349. https://doi.org/10.3390/info10110349

Viola J, Chen Y, Wang J (2021). FaultFace: Deep convolutional generative adversarial network (DCGAN) based ball-bearing failure detection method. Information Sciences 542: 195-211. https://doi.org/10.1016/j.ins.2020.06.060

Wang N, Nie J, Li J, Wang K, Ling S (2022). A compression strategy to accelerate LSTM meta-learning on FPGA. Information & Communications Technology Express 8 (3): 322-327. https://doi.org/10.1016/j.icte.2022.03.014

Xu P, Du R, Zhang Z (2019). Predicting pipeline leakage in petrochemical system through GAN and LSTM. Knowledge-Based Systems 175: 50-61. https://doi.org/10.1016/j.knosys.2019.03.013

Yang Y, Li Y, Yang J, Wen J (2022). Dissimilarity-based active learning for embedded weed identification. Turkish Journal of Agriculture and Forestry 46 (3): 390-401. 10.55730/1300-011X.3011

Yoon J, Jordon J, Schaar M (2018). Gain: Missing data imputation using generative adversarial nets. In: International conference on machine learning; pp. 5689-5698.

Zhang H, Goodfellow I, Metaxas D, Odena A (2019). Self-attention generative adversarial networks. In: International conference on machine learning; pp. 7354-7363.

Zhang L, Liu Y, Ren L, Teuling AJ, Zhang X et al. (2021). Reconstruction of ESA CCI satellite-derived soil moisture using an artificial neural network technology. Science of the Total Environment 782: 146602. https://doi.org/10.1016/j.scitotenv.2021.146602