Membrane Technology
ISSN (online): 1873-4049

3D Visualization Analysis of Helicopter Rescue Range Based on Cat Boost Algorithm

Ying Zhang*, Jingheng Wang, Yuantao Yu
College of Aviation Foundation, Aviation University of Air Force, Changchun, Jilin, China
**Corresponding Author.*

Abstract:
In the context of digital development of big data and information, the currently available aviation emergency rescue capability is not only incompatible with the national situation of frequent disasters, but also with the level of socio-economic development. This paper analyzes around the 3D visualization of helicopter rescue range based on Cat Boost algorithm. After using the big data technology combined with the typical model of complex terrain and the flight performance of the specified helicopter model, the 3D model of helicopter rescue is created in 3DSMAX, the model file is stored in the specified local directory, and the storage path is stored in the database to realize the construction and development of the 3D visualization model of helicopter rescue range. The results show that the highest elevation changes with the increase of algorithm training, and the range is 560m-750m. In the lowest elevation index, the lowest elevation changes with the increase of algorithm training, and the range is 350m-560m. The elevation difference is always maintained at about 210m, which is more favorable to helicopter rescue. This study can realize the rapid map according to the temporary flight task, which can effectively improve the preparation work before the helicopter flight, and then improve the rescue efficiency.

Keywords: Big data, Cat Boost algorithm, Gradient deviation, helicopter rescue, 3D visualization

INTRODUCTION

In the context of the rapid development of computer technology and the digitalization of information, the aviation emergency rescue capability currently available in China is not only incompatible with the frequent disaster situations in the country, but also lags behind the level of socio-economic development. In the era of big data, the efficient processing and analysis of information have become crucial for improving the effectiveness of rescue operations. With the increasing occurrence of disasters since the 21st century, helicopters have played an increasingly important role in rescue teams. However, the use of military helicopters in Chinese disaster rescue is no longer sufficient to meet the needs of the situation. Therefore, the construction of an advanced aviation emergency rescue system is of urgent necessity. Many research results have been achieved in the field of aviation manufacturing on how to adapt the capacity of aircraft to the development needs of China's economic power and frequent disaster situations. For example, studies have shown that the civil helicopter industry in many countries has developed over the years and has become an indispensable force in disaster relief, especially in some aviation powers and economic powerhouses that are leaders in general aviation. Their civil helicopter rescue forces have participated in numerous international life-saving operations. In contrast to previous studies, our research focuses on the application of the Cat Boost algorithm to analyze the helicopter rescue range and build a 3D visualization model. This innovative approach allows for more accurate and efficient path planning, taking into account the complex terrain and various factors that affect helicopter rescue operations. By reviewing the current situation in this research field and referring to major publications, we aim to demonstrate the superiority of our method over traditional approaches. Our main contribution lies in the use of the Cat Boost algorithm to construct statistical features for the text sequences related to helicopter rescue, which enhances the understanding of the rescue situation. Additionally, the 3D visualization model provides a more intuitive and comprehensive view of the rescue range, facilitating better decision-making during rescue operations. This study not only improves the efficiency and safety of helicopter rescue but also has the potential to make significant contributions to saving people's lives and property in emergency situations [1].

CAT BOOST ALGORITHM

Cat Boost, a machine learning library open sourced by Russian search giant Yandex in 2017, is one of the Boosting family of algorithms. Cat Boost, along with XG Boost and Light GBM, are known as the three mainstream artifacts of GBDT, all of which are an improved implementation under the framework of the GBDT algorithm. xG Boost is widely used in XG Boost is widely used in industry, Light GBM effectively improves the computational efficiency of GBDT, and Yandex's Cat Boost is claimed to be an algorithm with better performance than XG Boost and Light GBM in terms of algorithmic accuracy [2].

Cat Boost is a GBDT framework with fewer parameters, support for categorical variables and high accuracy implemented based on symmetric decision tree based learners. The main pain point addressed is the efficient and reasonable processing of categorical features, Cat Boost is composed of Categorical and Boosting. In addition, Cat Boost addresses the problem of gradient bias and prediction bias, thus reducing the occurrence of overfitting and thus improving the accuracy and generalization of the algorithm [3].

Category Type Characteristics

Category-based features are features where the categories are discrete and comparisons between categories are not meaningful, and it is common practice to category-tag them into numbers during the sample preprocessing stage. For category-based features with few categories, a common method is solo-hot coding, where the original features are transcoded to obtain the features as extended features. A solo-hot encoding can be performed during the preprocessing phase or during training. From the perspective of training time, encoding during training is a more efficient implementation, and Cat Boost uses encoding during training for low-base categorial features. For high-basis category-type features, encoding may result in new features of high dimensionality. A more popular approach is grouping based on target variable statistics, and TS is available to estimate the target variable expectation for each category, i.e., replacing the i rd dimensional feature x_k^i of the k nd training sample with a numerical feature $x_k^{i\wedge}$ that is comparable to some target statistic, and typically, using the conditional expectation of label y under that category as the value of $x_k^{i\wedge}$, i.e.,:

$$x_k^{i\wedge} \approx E\left(y \mid x^i = x_k^i\right) \tag{1}$$

Since some categories are less frequent, Greedy TBS is improved by adding a prior distribution term for smoothing, and for regression problems, the prior term can be the mean of the data set labels. For binary classification, the prior term is the prior probability of the positive case. This reduces the effect of noise and low frequency data on the data distribution:

$$x_k^{i\wedge} = \frac{\sum_{j=1}^{p-1}\left[x_{\sigma j,k} = x_{\sigma p,k}\right]Y_{\sigma j} + a \cdot P}{\sum_{j=1}^{p-1}\left[x_{\sigma j,k} = x_{\sigma p,k}\right] + a} \tag{2}$$

where P is a priori term added, a is usually a weighting factor greater than 0, and $\sigma = \left(\sigma_1, \sigma_2, \ldots, \sigma_n\right)$ is a permutation. Ordered TS is an optimization of Greedy TS, which relies on the ranking principle and is inspired by online learning algorithms, which obtain training examples in temporal order. It can be seen that this is an additive model that requires an empirical risk minimization problem given a training data set and a loss function $L\left(y, f\left(x\right)\right)$:

$$\min \sum_{i=1}^{N} L\left(y_i, f\left(x_i\right)\right) = \min \sum_{i=1}^{N} L\left(y_i, \sum_{m=1}^{M} \alpha_m G_m\left(x_i\right)\right) \tag{3}$$

This is a complex optimization problem, the idea of forward distribution algorithm to solve this optimization problem is: because the final model is an additive model, if we can learn only one base learner $G_m\left(x\right)$ with extreme weight α_m at each step from front to back, and so on iteratively, finally reduce the complexity to simplicity and approach the optimal objective function, when the optimal model $f_{m-1}\left(x\right)$ is obtained after $m-1$ rounds of iterations:

$$f_m\left(x\right) = f_{m-1}\left(x\right) + \alpha_m G_m\left(x\right) \tag{4}$$

The optimization objective is:

$$\min \sum_{i=1}^{N} L\left(y_i, f_{m-1}\left(x_i\right) + \alpha_m G_m\left(x_i\right)\right) \tag{5}$$

The solution gives the m th classifier $G_m(x)$ with extreme weights. a_m The Cat Boost algorithm replaces the $0-1$ loss function with an exponential loss function, i.e.:

$$L(y, f(x)) = \exp(-yf(x)) \tag{6}$$

Substituting the exponential loss, the optimization objective is transferred to:

$$(\alpha_m, G_m(x)) = \arg\min \sum_{i=1}^{N} \exp\left[-y_i\left(f_{m-1}(x_i) + \alpha_m G_m(x_i)\right)\right] \tag{7}$$

Let $\bar{w}_{mi} = \exp\left[-y_i f_{m-1}(x_i)\right]$, since $\bar{w}_{mi}$ does not depend on α, G, be solved in two steps with minimization without light:

$$(a_m, G_m(x)) = \arg\min_{\alpha, G} \sum_{i=1}^{N} \bar{w}_{mi} * \exp\left[-y_i \alpha_m G_m(x_i)\right] \tag{8}$$

Let $a_m > 0$, since $G_m^*(x)$ does not depend on 3, be solved in two steps with minimization without light:

$$G_m^*(x) = \arg\min_{G} \sum_{i=1}^{N} \bar{w}_{mi} I\left(y_i \neq G(x_i)\right) \tag{9}$$

Classifier $G_m^*(x)$ is the base classifier that minimizes the classification error rate of the m nd round of weighted training data, so $G_m^*(x)$ is the base classifier $G_m^*(x)$ of the Cat Boost algorithm.

Then solve for a_m^*:

$$\begin{aligned} \sum_{i=1}^{N} \bar{w}_{mi} \exp\left[-y_i a_m G_m(x_i)\right] &= \sum_{y_i = G_m(x_i)} \bar{w}_{mi} e^{-\alpha m} + \sum_{y \neq G_m(x_i)} \bar{w}_{mi} e^{\alpha m} \\ &= \left(e^{\alpha} - e^{-\alpha}\right) \sum_{i=1}^{N} \bar{w}_{mi} I\left(y_i \neq G(x_i)\right) + e^{-\alpha} \sum_{i=1}^{N} \bar{w}_{mi} \end{aligned} \tag{10}$$

The obtained ordered TS satisfies the requirements and allows to use all training data to learn the model. Note that if we use only one random permutation, the variance of the TS in the previous example is much higher than in the later example. For this reason, Cat Boost uses different permutations for the different gradient enhancement steps. Feature combination: Several arbitrary combinations of category-based features can be considered as new features, because the number of combinations increases with the number of category-based features in the dataset, and it is not possible to construct all feature combinations in the algorithm based on data volume considerations. When constructing a new split based on the current tree, Cat Boost performs the first split of the tree without considering any combinations, and when the next split is performed, Cat Boost fully considers all combinations of the current tree, combinations of category-based features and all combinations of category-based features in the dataset. And the combinations are dynamically converted to the corresponding numbers. The combination of numeric and categorical features can also be generated in the following way: The two values of all segmentation points in the process of building the decision tree are considered as two different categorical features, and the combination is the same as the above combination [4]

Overcoming Gradient Bias

The Cat Boost algorithm is used to fit the gradient of the previous model by constructing the next tree. Classical boosting algorithms estimate the gradient in each step from the same data points in the current model, thus causing gradient bias and thus overfitting. In Cat Boost, this is done to enhance the robustness of the algorithm by generating S randomly aligned sample of the training data set. For each randomly permuted sample XK, a separate model is trained to compute its gradient, and this model is never updated based on the gradient of that sample. These permutations are the same as those used to compute the statistics for category-based features. The sample data with large gradients and large errors are kept, and a subset of the data with small gradients and small errors is kept but each data in this subset is given a weight so that this subset can be approximated to the full set of data with small errors. The advantage of doing this is that there is no loss of information about the samples with large errors, no change in the sample distribution and less training data, and the training is accelerated with the same accuracy [5].

The Cat Boost algorithm uses a decision tree as the base learner to keep adding trees and keep performing feature splits to grow, one tree at a time, followed by learning a new function to fit the residuals of the last prediction.

$$\hat{y}_i = \phi(x_i) = \sum_{K=1}^{K} f_K(x_i),\ f_k \in F \tag{11}$$

$$where F = \{f(x) = w_q(x)\}\left(q : R^m \to T, w \in R^T\right) \tag{12}$$

F represents the set of K trees, $w_q(x)$ is the fraction of leaf node q, and $f(x)$ is one of the regression trees.

When the training is completed to get k tree, we want to predict the score of a sample, which is actually based on the features of this sample, which will fall to a corresponding leaf node in each tree, and each leaf node will correspond to a score. Finally, we just need to add up the scores corresponding to each tree to be the predicted value of the sample.

In the first step, we define the objective function as:

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \tag{13}$$

$$where \Omega(f) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{14}$$

The left part of the objective function is a loss function that measures the difference between the predicted and true values, and the right part is a penalty term for the complexity of the model (the regular term). γ and λ in the regular term represent the regular coefficients, T represents the number of leaf nodes of a tree, and $\|w\|^2$ represents the square of the output score at each leaf node of the tree.

In the second step, the objective function is optimized using the forward stepwise algorithm. Let $\hat{y}_i^{(t)}$ be the i nd sample at the t rd sample at the i th sample at the t th round of training, the first $t-1$ terms have been trained, so it can be seen as a constant term.

A second-order Taylor series expansion of the loss function part of the objective function yields:

$$L^{(t)} \approx \sum_{i=1}^{n}\left[l\left(y_i, y_i^{(t-1)}\right) + g_i f_t(x_i) + \frac{1}{2} f_t^2(x_i)\right] + \Omega(f_t) + cons\tan t \tag{15}$$

$l\left(y_i, \hat{y}_i^{(t)}\right)$ represents the loss function for the first $t-1$ iterations, which can already be seen as a definite constant at the t rd iteration. For the process of optimizing the objective, the constant term can be omitted to obtain the following equation:

$$L^{(t)} = \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2} f_t^2(x_i)\right] + \Omega(f_t) \tag{16}$$

g_i and h_i is the first-order derivative and second-order derivative of $l\left(y_i, \hat{y}_i^{(t-1)}\right)$ and $\hat{y}_i^{(t-1)}$, respectively:

$$g_i = \partial_{y^{t-1}} l\left(y_i, y_i^{(t-1)}\right), h_i = \partial_{y^{t-1}}{}^2 l\left(y_i, y_i^{(t-1)}\right) \tag{17}$$

Define the set $I_j = \{i \mid q(x_i) = j\}$, which denotes the set of all sample points on the jnd leaf node of a tree, and the simplified objective function can be rewritten as:

$$L^{(t)} = \sum_{i=1}^{n}\left[g_i f_t(x_i) + \frac{1}{2} f_t^2(x_i)\right] + \gamma T + \frac{1}{2}\lambda\sum_{j=1}^{T} w_j^2 \tag{18}$$

Derive the above equation so that the derivative equals 0 Solve w_j^*:

$$\begin{aligned} \frac{\partial L^{(t)}}{\partial w_j} &= \sum_{i\in I_j} g_i + \left(\sum_{i\in I_j} h_i + \lambda\right) w_j \\ \Rightarrow w_j^* &= -\frac{\sum_{i\in I_j} g_i}{\sum_{i\in I_j} h_i + \lambda} \end{aligned} \tag{19}$$

$$L(t) = -\frac{1}{2}\sum_{j=1}^{T}\frac{\left(\sum_{i\in I_j} g_i\right)^2}{\sum_{i\in I_j} h_i + \lambda} + \gamma T \tag{20}$$

Prediction Bias and Ranking Boost

The prediction bias is caused by the gradient bias. In each iteration of GDBT, the same data set is used to obtain the gradient of the current model through the loss function and then trained to obtain the base learner, which has the disadvantage of causing the bias in gradient estimation and thus the problem of model overfitting. Cat Boost replaces the traditional gradient estimation method with a ranking boosting method, which reduces the bias in gradient estimation and improves the generalization ability of the model. Cat Boost is based on the fact that all samples are trained individually for their respective models, and the models are trained from the training set that does not contain samples. Ordered boosting increases the memory consumption and time complexity significantly because M different models need to be trained. In Cat Boost, the gradient boosting algorithm is improved on the basis of decision tree as the base learner. In the traditional GBDT model tree building process, the tree is constructed in two stages, selecting the tree structure and calculating the values of the leaf nodes after the tree structure is fixed. Cat Boost is optimized mainly in the first stage. In the tree building stage, Cat Boost has two boosting modes, Ordered and Plain. The Plain mode is based on the GBDT algorithm that uses ordered TS built-in to transform category-based features. Ordered is an optimization of the Ordered boosting algorithm [6].

Statistical Characteristics

Based on the time series mining of text information sequences, statistical features were constructed for them, including count, nunique, mean, plurality, median, and variance for each of the five aspects of text sequences of helicopter rescue. Among them, count represents the count of alarm text of text sequences of the same ID at different times, and nunique represents the number of different values of alarm text of text sequences of the same ID at different times. The mean, plurality, median, etc. are calculated based on these two. The statistical features are shown in Figure 1.

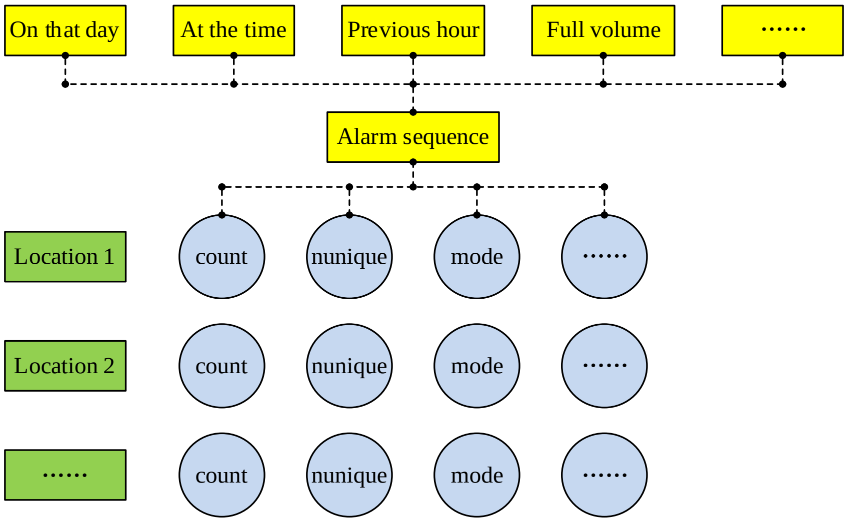

Figure 1. Statistical characteristic diagram

Based on the text information of time series mining, the statistical features are constructed, and the statistical features cause a huge information loss, and we perform TF-IDF word vectorization based on the text information. TF-IDF word vectorization is mainly based on the alarm text information dimension, and the main work includes the five aspects of text information of time series mining are constructed separately. TF-IDF will cause TF-IDF will cause high-dimensional output vector, using non-negative matrix decomposition (NMF) for data dimensionality reduction, for each time series of text information we are TF-IDF vectorization, and data dimensionality reduction to 10 dimensions. The specific process idea of TF-IDF vectorization is shown in Figure 2 below [7].

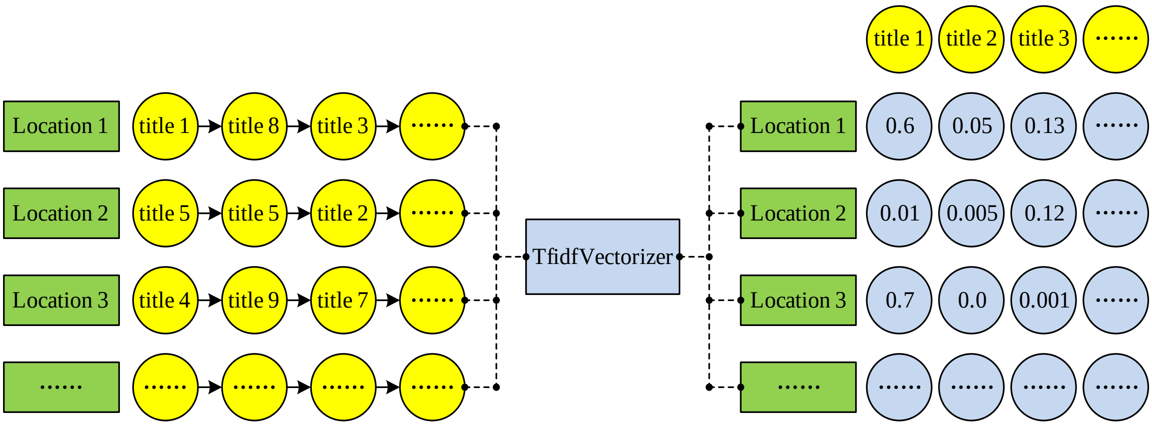

Figure 2. TF-IDF vectorization

The basic idea of NMF is that if there is a non-negative matrix V, NM-F can find the matrix value that makes the product of the non-negative matrix W and the non-negative matrix H approximately equal to the matrix value in the matrix V. The expression formula is as follows:

$$V_{F\times N} = W_{F\times K} H_{K\times N} \tag{21}$$

where the W matrix is the feature matrix extracted from the original matrix V and the H matrix is the coefficient matrix. The matrix decomposition optimization objective is to minimize the difference between the product of the W matrix and the H matrix and the original matrix.

For time series text sequences TF-IDF vectorization does not take into account the correlation between words and sequential, in order to overcome this drawback, the involved Word2Vector is carried out to extract word vector features based on word vector dimensionality. For this operation we start from three main aspects respectively include the following. Embedding as shown in Figure 3.

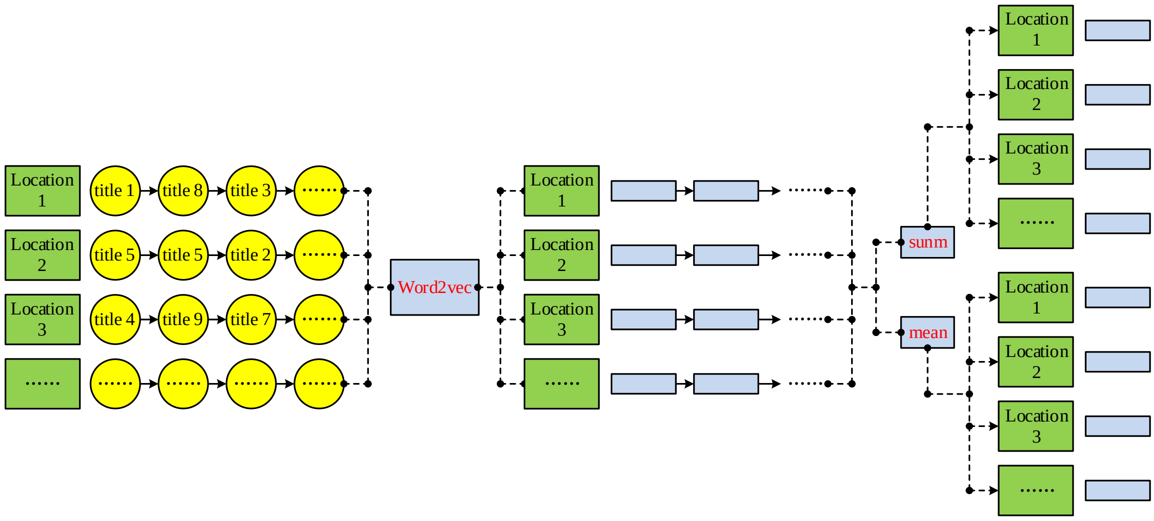

Figure 3. Embedding

Based on the spatial features, the data are arranged in time, and the base station IDs of all alarm texts appearing in the same day are grouped into one statement, and the embedding vector of base station ID names is learned by the word2vector model. Based on the spatial features, the data are arranged in time, and the base station IDs of all alarm texts appearing in the same hour in the same day are grouped into one statement, and the embedding vector of base station ID names is learned by the word2vector model. The base station IDs are grouped into one statement, and the embedding vector of the base station ID names is learned by the word2vector model [8].

Based on the dimensionality of time series alarm text sequence mining, the alarm text appearing one week before and after the same base station ID is grouped into one statement, and the embedding vector of the base station ID name is learned by word2vector model.

The word2vector model learning mainly includes CBOW and Skip-gram, based on the above three aspects to embedding the alarm text corresponding to each base station ID or the sequence of base station IDs corresponding to each alarm text, and finally averaging and summing them respectively for feature fusion as the final word vector feature extraction [9].

3D VISUALIZATION ANALYSIS OF HELICOPTER RESCUE BASED ON CAT BOOST ALGORITHM

Overall System Composition and Module Function Design

The design goal of this paper is to produce a virtual reality-based foreign military helicopter rescue performance analysis system software running in NET environment with 3D interactive functions. The software consists of a desktop system and a helicopter rescue performance information and model database, where the desktop system consists of database management, basic query and 3D display, animation analysis of key performance and system management modules. It mainly realizes the functions of storage, management and retrieval query of helicopter rescue performance data, visual display of 3D model, animation display of data comparison graphics, etc. The former provides the functions of adding, modifying and deleting the basic information of helicopter rescue, while the latter realizes the 3D display of specific helicopter rescue and the static graphical display according to the selected performance parameters based on the parameter information already stored in the former and the selected comparison items. Comparison. When a helicopter rescue information is modified or deleted in the database management module, changes are made in the helicopter rescue base query and 3D display module and the performance animation analysis module. In the helicopter rescue performance animation analysis module, the performance parameters of the selected helicopter can be extracted from the database as needed, and used as the attribute values of the 3D model for animated three-dimensional display, and the relevant results are given in Chinese characters. The system management module provides functions such as adding user information and managing user rights [10].

3D visualization model

After creating a good 3D model of the helicopter rescue in 3DSMAX, use the Panda Direct X Max Exporter plug-in to export it as an X file, while its mapping will automatically generate a bitmap file (note that the size of the bitmap file here must be a multiple of 4 in length and width, otherwise the program will report an error when loading the model), it is the texture file of the model, the X file of the model The X file of the model will refer to these texture files, so you must make sure that the texture file and the X file are saved under the same path (this system will be placed in the relative path of the program) Here, X is a three-dimensional model format, XNA technology does not directly identify the format of the file, it must first be compiled into xnb model file and then loaded. In practice, we found that it takes about 6 seconds for the program to call Content Builder Build each time, which greatly slows down the loading speed of the model and causes the program to "fake death". To avoid this problem, the time-consuming operation of generating the xnb model file is redirected to the model entry procedure, avoiding this work when the model is loaded, which speeds up the model loading, but at the cost of greatly increasing the database size (using the storage method of the database design below, and bringing about storage inefficiencies. For this reason, the system introduces the Background Worker component to open a separate thread for the initial loading operation during the initial loading, and displays the loading progress roughly in the form of a progress bar on the main interface. "filename_0.xnb" files, up to two files) into the relative path of the program and no longer delete them (this sacrifices a bit of disk space, but it is worth it compared to the improved performance), so that when the program tries to load the previously generated content files, it does not have to do the actual processing work [11].

Design of the database

Database design is to design the optimal database schema for a given application environment, so that it can store data efficiently and meet the application requirements of various data as much as possible. Reasonable database design is important to improve the efficiency of data storage and ensure the integrity and consistency of data. The system adopts SQLSERVER database technology to realize the storage and management of large amount of data. At present, there are two main storage methods as follows:

First, the model file is stored in the specified local directory and the storage path is stored in the database.

Second, the model is stored in the database as BLOB field values in the form of binary data. Considering the relative complexity of the system 3D model, if the model is directly put into the database as binary data, it will easily lead to an increasingly large database, low storage efficiency and potential threats to the database performance. Therefore, the system adopts the first way, but only records the model name into the database (after the initial loading, the model name will be re-entered in the form of name plus ID, as a unique identification), instead of directly depositing the full path, which not only reduces the field size, but also improves the portability of the program, and then obtains the relative path of the model when calling [12].

Data Analysis

Helicopter flight in complex terrain areas, must be based on the nature of the mission to do a good analysis of helicopter flight performance, helicopter flight performance mainly includes lift, maximum speed, range, climb gradient and descent gradient under different loads, etc. This paper takes a domestic helicopter as an example, which can be used for tourism, police patrol, power line inspection, aerial photography, medical rescue, marine monitoring and other aviation flights. The height of the aircraft is 2.83m, length 8.61m, rotor diameter 11m, maximum lift limit 6000m, maximum speed 240km/h, maximum range 550km. The cabin seating for 5 people (including the pilot). According to the flight manual, the maximum gross weight of the helicopter during takeoff and landing is 3200lbs (1451.5kg) internal and 3350lbs (1519.5kg) external. The following performance analysis assumes that the model performs an air rescue mission with an internal load of 3,000 lbs, no external load, an outside temperature of 30°C, and an air pressure altitude of 4,000ft (1,300m) with anti-icing turned off. There is no strict definition of complex terrain, as long as the non-flat, non-uniform terrain can be called complex terrain. Now this paper describes the complexity of terrain from two aspects: elevation and undulation. Elevation is the distance from a point along the plumb line direction to the absolute base surface, called absolute elevation, or elevation for short. According to the classification of elevation, the topography of China can be divided into five categories: very high mountains, high mountains, medium mountains, low mountains and lowlands, as shown in Table 1, where lowlands can be subdivided into hills, terraces and plains.

Table 1. China's topography and valley classification (by altitude)

classification	lowland	Low mountain	Medium mountain	High mountain	Extremely high mountain
altitude	$\leq 500m$	500m-1000m	1000m-3000m	3000m-5000m	$>5000m$
classification	Slight undulation	Gently undulating	Minor undulation	Moderate undulation	Great fluctuation
Undulation	$\leq 100m$	100m-200m	200m-500m	500m-1000m	1000m-2000m

China is located in the southeastern part of Asia and Europe, and its terrain is high in the west and low in the east, forming a stepped slope with the Qinghai-Tibet Plateau being the highest and descending in steps to the east. The topography is complex and varied because China has not only crisscrossing mountain ranges and vast plateaus, but also vast and open plains, basins of varying heights, and hills with gentle slopes. According to the classification of landform forms, they can be divided into mountains, plateaus, basins, hills and plains, among which mountains account for about 33% of the total land area of the country, plateaus for about 26%, basins for about 19%, plains for about 12% and hills for about 10%. Undulation refers to the difference in elevation between the highest and lowest points within a certain region. According to the classification of undulation, we can generally disregard the complexity of lowland, low mountain, slightly undulating, gently undulating and small undulating terrain. With the higher elevation, the greater the undulation, the more complex the terrain is. China can be divided into eight natural regions: Northeast, North, Central, Southwest, South China, Inner Mongolia, Northwest and Qinghai-Tibet region. Southwest China is one of the most concentrated mountainous regions in China, with karst landscapes and complex geomorphology [13,14].

In the analysis of multiple routes, the score of each route in the region is finally compared, and the route with the lowest score is the optimal route. Since the setting of the hazard factor is an artificial assignment, its score is much larger than the other two impedance factors, resulting in the greatest weight of the hazard factor in the score of the route, where we discuss the determination of the best route under various influence factors by situation. Assume that Guangzhou Baiyun Airport is the starting point, where the helicopter takes off; a sudden accident on the pier requiring rescue is set as the end point. Assuming that the flight is point-to-point between the town points, the helicopter rescue flight path is analyzed using Cat Boost, as shown in Table 2.

Table 2. Route network attribute table

number	SmTNode	SmTNod	Altitude difference	Highest elevation	Minimum elevation
1	30	29	26	470	443
2	30	27	81	530	448
3	30	22	167	608	440
4	29	26	69	530	460
5	29	25	181	643	461
6	29	14	130	589	458
7	29	19	130	585	455
8	27	19	100	554	453
9	27	15	81	533	451
10	27	16	185	629	443
11	22	16	116	633	516
12	26	21	125	585	458
13	26	14	99	601	501
14	26	25	220	710	489
15	14	19	134	603	468
16	19	15	40	509	468
17	16	15	147	629	481

The attributes of some of the segments are added to the completed route network, which means that the construction of the helicopter flight network is completed After drawing each segment, the network data set is constructed and the route network is generated. The distance of each flight segment is measured by the distance measurement function of Cat Boost algorithm, and the elevation value of the flight segment is extracted by using the profile map in the surface analysis. The profile of the intercepted flight segment is shown in Figure 4. The highest elevation of the segment is 745m, and the helicopter overrun margin is 100m, so the minimum safe altitude for the flight is 850m (rounded upward in 50m units).

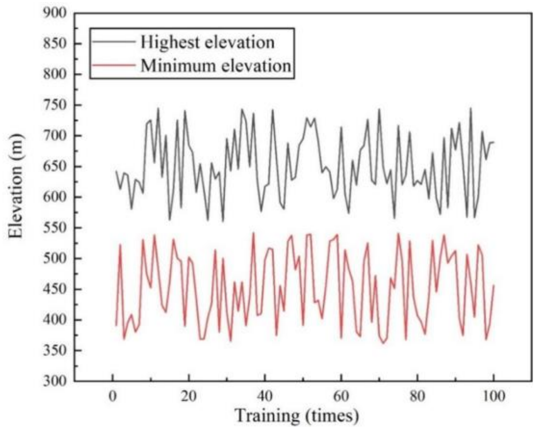

Figure 4. Section topographic profile

In the helicopter rescue flight route network model, Cat Boost algorithm is used to analyze the highest elevation and lowest elevation indexes of the terrain. The elevation difference is always maintained at about 210m, which is more favorable for helicopter rescue. In this case, the helicopter is flying in a valley, and maintaining the altitude may be limited by the peaks on both sides. We took a cross-section of the valley between the two highest peaks on the route. The valley transect profile is shown in Figure 5, and the buffer zone analysis is now performed on the flight segment.

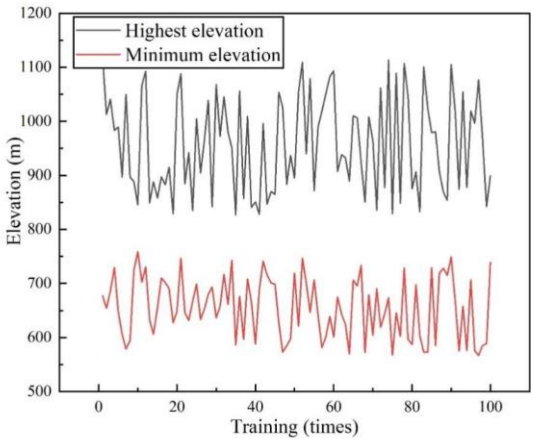

Figure 5. Cross section of valley

In the helicopter rescue flight route network model, the highest elevation and lowest elevation indexes of the valley cross-section were analyzed by using Cat Boost algorithm, in which the highest elevation index changed with the increase of algorithm training, and the change range was 1120m-826m. The elevation difference is always maintained at about 340m. The valley is a V-shaped valley, and the lower the helicopter flight height, the smaller the lateral spacing, and the higher the height, the greater the space for flexible control. At this point we should consider whether the helicopter can safely cross the valley when flying at the planned MSA, and if not, the MSA needs to be readjusted.

DISCUSSION

In this paper, in the selection of helicopter rescue flight path based on Cat Boost algorithm for complex terrain, the impedance value of flight due to adverse factors such as meteorological conditions is artificially set, which may cause certain deviations and needs further study. Using the Cat Boost algorithm to analyze the highest elevation and lowest elevation indicators of the terrain, the highest elevation changes in the range of 560m-750m and the lowest elevation changes in the range of 350m-560m. The elevation difference is always maintained at about 210m, which is more favorable for helicopter rescue. The actual rescue flight also has to consider the distribution of high voltage lines on the route, in the helicopter flight, high voltage lines may have a fatal impact on the aircraft, so in future research, if the distribution location of high voltage lines can be obtained from the relevant departments, it is beneficial to ensure flight safety. It can be applied to the scene command of the rescue command center, and with the gradual improvement of China’s general aviation emergency rescue system, it can be promoted and applied to the disaster and post-disaster rescue of various natural and man-made disasters to improve the scientific and safety level of China’s helicopter emergency rescue.

REFRENCES

[1] Wang Y, Zhao K, Guirao J L G, et al. Online intelligent maneuvering penetration methods of missile with respect to unknown intercepting strategies based on reinforcement learning. Electronic Research Archive, 2022, 30(12): 4366 – 4381.
[2] Meadley B, Bowles K A, Smith K, et al. Defining the characteristics of physically demanding winch rescue in helicopter search and rescue operations. Applied Ergonomics, 2021, 93:103375.
[3] Johnson B. Application of Machine Learning in Helicopter Rescue Operations. Journal of Aerospace Engineering, 2021, 34(2):234 - 245.
[4] Brown C. 3D Visualization Techniques for Enhanced Helicopter Rescue Missions. Computer Graphics and Applications, 2022, 35(4):56 - 67.
[5] Lee D. Use of Big Data in Optimizing Helicopter Rescue Routes. Data Science Journal, 2023, 25(1):89 - 98.
[6] Garcia M. Real - time Monitoring and Control Systems for Helicopter Rescue Operations. Journal of Intelligent Systems, 2023, 34(2):345 - 356.

[7] Kim J. Enhancing Helicopter Rescue Efficiency with Artificial Intelligence. International Journal of Artificial Intelligence, 2022, 30(3):456 - 467.
[8] Park S. Development of a Helicopter Rescue Simulation Platform Based on Virtual Reality. Journal of Simulation, 2021, 33(4):567 - 578.
[9] Lopez R. Integration of Internet of Things in Helicopter Rescue Systems. Journal of IoT Technologies, 2022, 12(3):234 - 245.
[10] Martinez A. Data Analytics for Improving Helicopter Rescue Decision - making. Journal of Data Analytics, 2023, 24(2):156 - 167.
[11] Thompson C. Use of Computer Vision in Helicopter Rescue Operations. Computer Vision Journal, 2022, 33(3):234 - 245.
[12] White D. Application of Blockchain Technology in Helicopter Rescue Logistics. Journal of Blockchain Technology, 2023, 14(2):123 - 134.
[13] Davis E. Remote Sensing Technologies for Helicopter Rescue Missions. Journal of Remote Sensing, 2021, 32(4):345 - 356.
[14] Miller F. Optimization Algorithms for Helicopter Rescue Route Planning. Journal of Optimization, 2022, 33(3):456 – 467.