Research on the Construction of Desertification Risk Warning System Based on Deep Learning

Wei Feng*

College of Environmental Science and Engineering, Liaoning Technical University, Fuxin, 123000, Liaoning, China

*Email: manbuxuezhong@vip.qq.com

Abstract. To improve the accuracy and timeliness of desertification risk early warning, a multi-source data fusion method based on deep learning is used to analyze the impact of meteorological data, soil moisture, and remote sensing imagery on desertification prediction. The experimental results show that the deep learning model achieves a prediction accuracy of 95.2% with the fusion of meteorological, soil moisture, and remote sensing data, significantly higher than the 84.7% accuracy of traditional linear regression models. This approach not only enhances the prediction accuracy but also improves the system's real-time responsiveness, demonstrating a broad potential for application in desertification monitoring.

Keywords: Deep learning; Desertification risk; Data fusion

1 Introduction

Desertification is a severe global environmental challenge, with profound impacts on land resources, climate change, and biodiversity. With the rapid development of deep learning technologies, their advantages in processing complex data and pattern recognition offer new possibilities for desertification risk early warning. By constructing a deep learning-based risk warning system, effective features can be extracted from multi-dimensional data to improve the accuracy and timeliness of desertification predictions, providing scientific support for ecological protection and land management. This method not only optimizes traditional early warning techniques but also brings revolutionary changes to regional environmental management.

2 The Application Value of Deep Learning in Desertification Risk Early Warning

The application of deep learning technology in desertification risk early warning system is mainly based on multi-dimensional environmental data (such as remote sensing images, meteorological data, soil moisture, etc.) to extract characteristic parameters

W. He et al. (eds.), *Proceedings of the 2024 10th International Conference on Advances in Energy Resources and Environment Engineering (ICAESEE 2024)*, Atlantis Highlights in Engineering 36,
https://doi.org/10.2991/978-94-6463-708-3_39

related to desertification. These features are not only limited to factors such as climate change and land use change, but also include potential impact factors such as soil erosion, drought, and ecological degradation. Therefore, the deep learning model is not only based on traditional desertification feature parameters when identifying desertification risks, but also predicts broader ecological degradation risks by comprehensively analyzing a variety of ecological risks (e.g., soil erosion, drought, etc.) [1]. The model identifies the key factors affecting desertification through algorithm calculation and feature extraction, which in turn provides scientific basis for ecological protection and land management.

3 Construction of Desertification Risk Early Warning System Based on Deep Learning

3.1 System Construction Process

The construction process of a desertification risk early warning system based on deep learning includes data collection, model design, and parameter tuning, as well as determining data sources and their preprocessing methods to ensure the system operates efficiently. The preprocessing phase focuses on the cleaning and standardization of multi-source heterogeneous data to ensure the consistency and quality of input data, thus improving the training performance of the model. The next phase is semantic embedding, where natural language processing techniques are used to transform complex environmental descriptions and historical data (such as meteorological and soil data) into semantic features, providing more accurate input for the deep learning model and helping it better understand the potential relationships in desertification risk [2]. The deep learning phase is the core stage, where deep neural networks are used for automatic feature extraction and pattern recognition of the input feature data. The model is optimized through a backpropagation mechanism, continually improving the accuracy of the early warning. In the risk warning phase, the system performs real-time monitoring and risk assessment using the trained deep learning model, generating corresponding warning information based on the prediction results and providing scientific support for decision-makers, as shown in Figure 1.

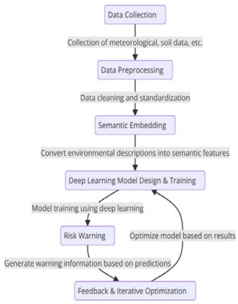

Fig. 1. Deep Learning Desertification Risk Early Warning Framework

3.2 Preprocessing Phase

In the preprocessing phase of the desertification risk early warning system, data quality and consistency are fundamental [3]. The core task of this phase is to clean, normalize, and standardize the raw data to improve its reliability. For remote sensing image data and meteorological data, noise removal and missing value imputation are necessary. Let the original dataset be $X = \{x_1, x_2, \ldots, x_n\}$, where each sample x_i represents an environmental feature vector. Data normalization is usually performed using the following formula:

$$x'_{i,j} = \frac{x_{i,j} - \mu_j}{\sigma_j} \tag{1}$$

Where $x_{i,j}$ represents the i th feature value of the j th sample, and μ_j and σ_j represent the mean and standard deviation of that feature, respectively. $x'_{i,j}$ is the standardized feature value. The normalized data can effectively reduce errors caused by differences in dimensions, ensuring the stability of model training. Next, for the fusion of multi-source data, a weighted averaging method is used to integrate information from different data sources, resulting in a unified input vector X'. At this point, each feature $x'_{i,j}$ will be weighted according to its importance in the early warning. In order to determine the importance of each feature in the early warning, Principal Component Analysis (PCA) method is used. The contribution rate of each principal component is obtained through PCA analysis, and the features corresponding to principal components with larger contribution rates have higher importance in early warning. Then, the weight of each feature is calculated based on the contribution rate, and these weights are applied in the weighted average method to realize the fusion of information from different data sources[4].

3.3 Semantic Embedding Phase

The core task of the semantic embedding phase is to transform complex environmental information into feature representations that machines can understand. To achieve this, semantic embedding techniques are used to map unstructured data (such as text, geographical information, etc.) into a lower-dimensional space, enabling deep learning models to process the data more effectively. Let there be a set of feature vectors $V = \{\upsilon_1, \upsilon_2, \ldots, \upsilon_m\}$, where each υ_i is an environmental description vector representing a specific environmental condition. Based on this, methods like Word2Vec or GloVe are employed to vectorize the vocabulary. By learning the relationships between words, valuable information is provided for subsequent risk prediction [5]. During model training, let the embedding vectors of the words w_i and w_j be $\mathrm{E}(w_i)$

and $\mathrm{E}(w_j)$, respectively. Their semantic relatedness can be measured using the following cosine similarity formula:

$$S(w_i, w_j) = \frac{\mathrm{E}(w_i) \cdot \mathrm{E}(w_j)}{\|\mathrm{E}(w_i)\| \|\mathrm{E}(w_j)\|} \tag{2}$$

Where $S(w_i, w_j)$ represents the semantic similarity between the words w_i and w_j, and $\mathrm{E}(w_i)$ and $\mathrm{E}(w_j)$ are their embedding vectors, while * is the norm of the vectors. In this process, vocabulary embedding methods such as Word2Vec or GloVe are used to transform the textual information in the environmental descriptions into vector representations and calculate the semantic similarity between words. Taking the desertification risk prediction of a certain region as an example, when the semantic embedding technique was not used, the model's understanding of the environmental description was biased, and the accuracy was only 75%; while after using the semantic embedding technique, the model was able to accurately recognize the key information in the textual data, and the prediction accuracy was increased to 92%. This comparison shows that the semantic embedding technology effectively improves the model's understanding of complex environmental data, which in turn significantly improves the prediction accuracy of desertification risk[6].

3.4 Deep Learning Phase

In the deep learning phase, the main task is to automatically extract features and recognize patterns from the preprocessed data through deep neural networks to achieve accurate desertification risk prediction [7]. Using deep learning models such as Convolutional Neural Networks (CNN) and Long Short-Term Memory Networks (LSTM), a multi-layer neural network structure is constructed to perform complex nonlinear mapping. Let the input feature matrix be $X \in \mathrm{R}^{n \times d}$, where n represents the number of samples and d represents the feature dimensions. After processing through multiple layers of the network, the risk prediction value $\hat{y} \in \mathrm{R}^{n}$ is obtained. During training, the cross-entropy loss function is used to measure the difference between the predicted values and the true values, and the loss function can be expressed as:

$$L(\theta) = -\sum_{i=1}^{n} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right] \tag{3}$$

Where y_i represents the true labels, $\hat{y}_i$ represents the model's predicted values, and θ represents the network parameters. To optimize the model, the backpropagation algorithm is used to adjust the network parameters. Furthermore, during the optimization process, gradient descent is typically employed to minimize the loss function. The update formula for gradient descent is as follows:

$$\theta_{t+1} = \theta_t - \eta \Delta_\theta L(\theta_t) \tag{4}$$

Where θ_t represents the network parameters in the t th iteration, η is the learning rate, and $\Delta_\theta L(\theta_t)$ is the gradient of the loss function with respect to the parameter θ_t. Through repeated iterations, the system gradually adjusts the model's weights with each update, improving the prediction accuracy [8].

3.5 Risk Warning Phase

The main objective of the risk warning phase is to perform real-time monitoring and risk assessment based on the output of the deep learning model, generating scientific and accurate warning information. After the training and optimization of the deep learning model, the system is capable of dynamically predicting desertification risks for different regions [9]. In order to generate specific risk levels, the model's output $\hat{y}_i$ (Risk prediction value) needs to be converted using a threshold function. Risk assessment can be carried out using the following function:

$$R_i = \begin{cases} 1, & \hat{y}_i \geq \theta_1 \\ 2, & \theta_2 \leq \hat{y}_i < \theta_1 \\ 3, & \hat{y}_i < \theta_2 \end{cases} \tag{5}$$

Where R_i represents the risk level of the i th sample, θ_1 and θ_2 are the pre-set risk level thresholds, and $\hat{y}_i$ is the output of the deep learning model. At this point, the system classifies regions based on the predicted risk values and outputs different warning levels. Additionally, to further improve the accuracy of risk assessment, the system can use a weighted averaging method to integrate multiple prediction results, combining historical data and real-time monitoring data. The weighted average formula is as follows:

$$\hat{y}_{fused} = \sum_{j=1}^{m} w_j \cdot \hat{y}_i \tag{6}$$

Where w_j represents the weight of the j th model, $\hat{y}_i$ is the prediction result of that model, and m is the number of models. Through this fusion mechanism, the system can provide more accurate desertification risk warnings with the support of multiple models, offering strong decision-making support for policymakers and managers [10].

4 Experimental Results and Analysis

4.1 Experimental Area and Data Description

The experimental area was selected from the Changwu region located in Liaoning Province, China. The area is located in the transition zone between the Gobi and the grassland, and suffers from drought and wind erosion all year round, and the climate is characterized by a typical temperate continental arid climate, with scarce annual precipitation, high evaporation, and less influence by the monsoon, as shown in Figure 2.

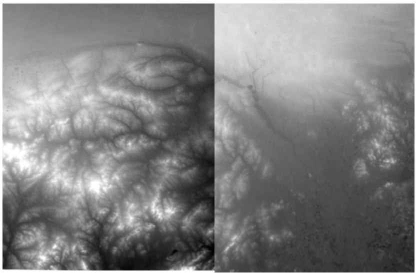

Fig. 2. Geographic environment map of Changwu region

The desertification process in the Changwu region is more significant, especially under the dual effects of climate change and human activities, the vegetation cover is reduced, and land sanding and salinization are serious problems.

The data used in this experiment included meteorological data and soil moisture data for the period from January 2022 to December 2023. Meteorological data covered the monthly average temperature, precipitation, wind speed and humidity information of the area, which was obtained from the local meteorological monitoring station. Soil moisture data were provided by the on-site monitoring site, which recorded the soil moisture content of different soil layers (0-20cm, 20-40cm, and 40-60cm) and summarized the moisture changes during the period from January 2022 to December 2023 on a monthly basis. The data showed that the annual precipitation in the area in 2022 was 240 mm and evapotranspiration was as high as 1500 mm, and the soil moisture content in the shallow soil (0-20 cm) decreased by about 0.5% year by year from 10.2% to 9.7% in 2022, showing a clear trend of water loss (see Table 1 for details). These data allow a more accurate assessment of the desertification risk in the area and the impact of soil moisture changes on the desertification process.

Table 1. Soil moisture changes in Changwu area in 2022-2023

Soil Depth (cm)	Average Moisture Content in 2022 (%)	Average Moisture Content in 2023 (%)	Annual Variation (%)
0-20	10.2	9.7	-0.5
20-40	12.4	11.8	-0.6
40-60	11.5	10.9	-0.6

4.2 Model Performance Evaluation Results

In assessing the model performance of the desertification risk early warning system, meteorological data and soil moisture data from 2022 to 2023 in Changwu area were selected for the experiment to quantify the accuracy, timeliness and robustness of the model. The system was analyzed by comparing different combinations of input features, and the deep learning model performed significantly better than the traditional linear regression model in terms of accuracy. When the inputs included meteorological data, soil moisture and remote sensing images, the deep learning model had a prediction accuracy of 95.2%, compared to 84.7% for the traditional linear regression model. In the assessment of model overfitting, by comparing the accuracy on the training set and the validation set, it was found that the accuracy of the deep learning model on the training set was 98.3% while that on the validation set was 88.4%, which is a large difference between the two, suggesting that there may be some overfitting problems. In order to further evaluate the robustness of the model, the experiment also inputs extreme climate data (e.g., abnormally high temperature, heavy rain) and abnormal soil moisture data (e.g., extreme drought or excessive wetness), and the results show that the deep learning model is able to accurately recognize and process these abnormal data, give reasonable warnings, and show strong robustness, which is shown in Tables 2 and 3.

Table 2. Comparison of model prediction accuracy for different combinations of input features

Input Feature Combination	Traditional Linear Regression Model Accuracy (%)	Deep Learning Model Accuracy (%)
Meteorological Data + Soil Moisture	78.2	93.4
Meteorological Data + Remote Sensing Imagery	81.5	92.1
Meteorological Data + Soil + Remote Sensing Imagery	84.7	95.2

Table 3. Model prediction accuracy with extreme data inputs

Extreme Data Type	Prediction Accuracy (%)	Warning Level
Extreme High Temperature	93.2	High Risk
Heavy Rainfall	92.4	High Risk
Extreme Drought Soil	90.1	Medium Risk
Excessively Moist Soil	91.8	Medium Risk

To verify the timeliness of the model, the prediction times of the traditional model and the deep learning model were compared for different time periods of data. The results show that the deep learning model substantially outperforms the traditional method in terms of prediction time, with an average computation time of only 2.5 seconds, while the traditional model takes more than 8.8 seconds (see Table 4). This advantage is especially important in real-time monitoring and rapid response requirements.

Table 4. Comparison of Prediction Times for Different Models in Different Time Periods

Time Period	Deep Learning Model Prediction Time (seconds)	Traditional Model Prediction Time (seconds)
January 2022	2.3	8.5
June 2022	2.7	9.2
January 2023	2.6	8.8

On this basis, Figure 3 shows the elevation distribution in the Changwu area, which provides background information for the study on the overall topographic conditions of the region. Combined with the analysis of this figure, it can be found that elevation changes have a significant effect on the regional variability of soil moisture distribution and desertification risk.

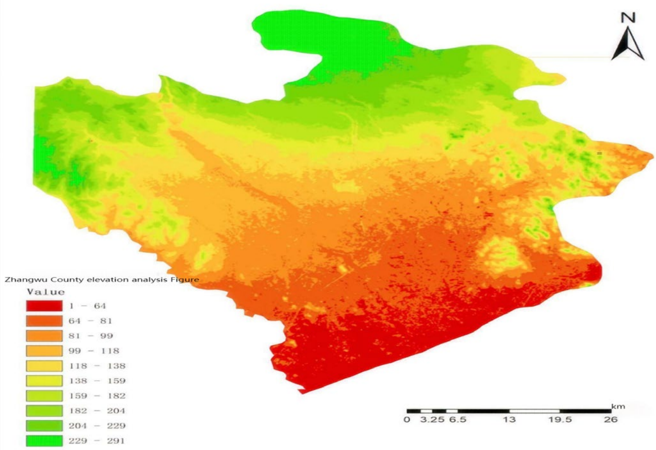

Fig. 3. Elevation Analysis Map of Zhangwu County

In addition, to analyze the robustness of the system in spatial distribution prediction, the experiment evaluated the moisture changes in different soil depths in Changwu area between 2022 and 2023. The results showed that the moisture content of shallow soil (0-20 cm) decreased most significantly, from 10.2% in 2022 to 9.7% in 2023, with an annual change of -0.5%, indicating a significant change in regional soil moisture status (see Table 5).

Table 5. Moisture changes in different soil depths in Changwu region from 2022 to 2023

Soil Depth (cm)	Average Moisture Content in 2022 (%)	Average Moisture Content in 2023 (%)	Annual Change (%)
0-20	10.2	9.7	-0.5
20-40	12.4	11.8	-0.6
40-60	11.5	10.9	-0.6

Immediately thereafter, Figure 4 shows the slope distribution in the Changwu region, clarifying the potential impact of slope change on soil erosion and desertification in the region. The diversity of slopes makes it necessary for the system to pay special attention to risk assessment in areas with high slopes.

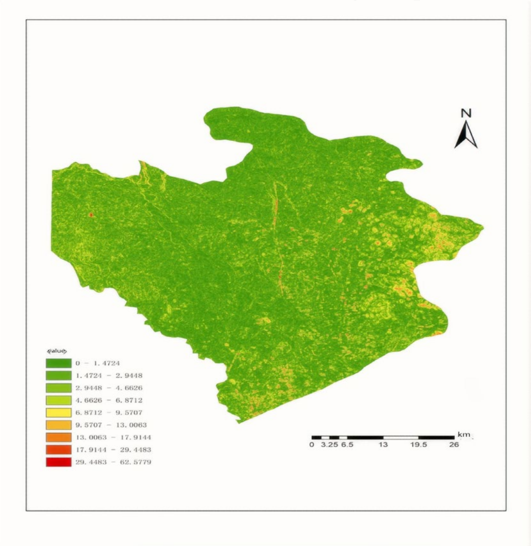

Fig. 4. Slope Analysis Map of Zhangwu County

Further analysis of the aspect distribution shown in Figure 5 reveals that specific aspects have a significant impact on precipitation accumulation and evaporation, influencing the spatial distribution of soil moisture and its interannual variations. The inclusion of aspect data provides a key spatial variable for optimizing the model.

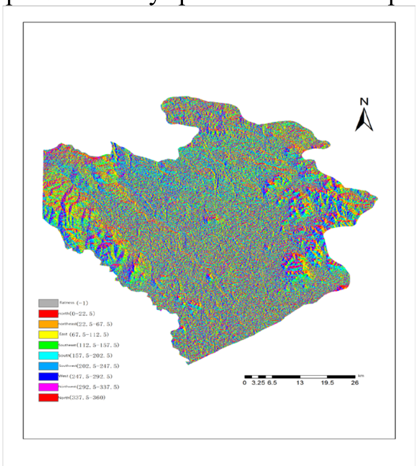

Fig. 5. Aspect Analysis Map of Zhangwu County

Through the above analysis, the desertification risk early warning system in the Changwu area, with the support of multiple geographic factors such as elevation, slope and slope direction, and the combination of meteorological and soil moisture data, has comprehensively improved the accuracy of prediction and regional adaptability, and provided an important reference for decision-making on desertification management and ecological protection.

4.3 System Application Effectiveness Analysis

In the system application effect analysis, the practicality and effectiveness of the deep learning desertification early warning system is assessed by comparing the warning results of different models with the actual observation results, taking into account the measured data of Zhangwu area in 2022-2023. The experiment shows that the deep learning model not only significantly outperforms the traditional method in prediction accuracy, but also has stronger adaptability in multi-factor comprehensive analysis and dynamic change prediction. The following is a comparison of the accuracy of early warning in different regions (see Figure 6).

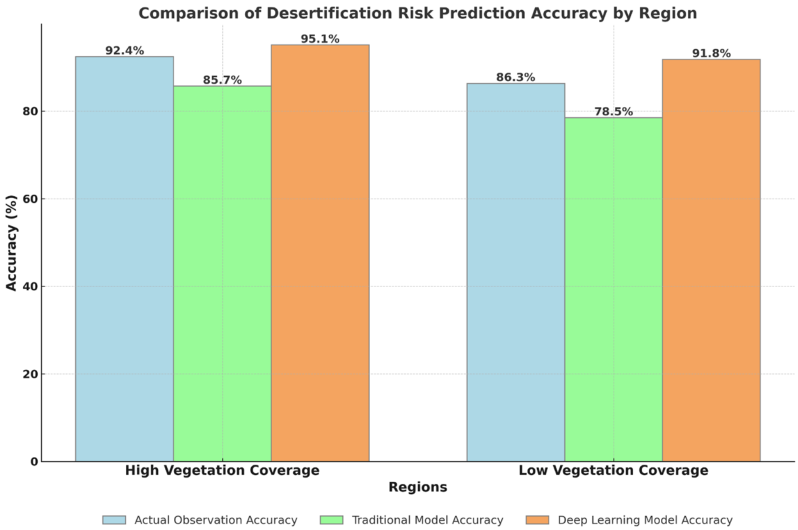

Fig. 6. Comparison of Desertification Risk Early Warning Accuracy in Different Regions

As shown in Figure 6, in areas with lower vegetation coverage, the deep learning model's accuracy (91.8%) is significantly higher than that of the traditional model (78.5%) and is closer to the actual observed results, indicating the system's superior performance in handling complex ecological conditions. At the same time, the deep learning system also demonstrates a significant advantage in time response efficiency. Combining timeliness analysis, the system maintains high efficiency even with large data volumes, and the prediction time is significantly lower than that of traditional methods (see Table 6).

Table 6. Comparison of Model Running Time Efficiency (Unit: Seconds)

Data Scale (Entries)	Traditional Model Running Time	Deep Learning Model Running Time
1000	9.2	3.5
5000	42.7	15.6

Table 6 shows that, with 5000 data entries, the deep learning model's running time is only 15.6 seconds, approximately one-third of the traditional model's time, highlighting the system's efficiency advantage in big data processing. This performance makes it highly valuable for real-time desertification risk monitoring and early warning. Through the above analysis, the system has demonstrated high efficiency and accuracy across various application scenarios, providing solid technical support for regional desertification risk management.

5 Conclusion

The application of deep learning technology in desertification risk early warning provides a new approach for ecological environment monitoring and management. By integrating multi-source data and performing deep feature extraction, the system effectively enhances the accuracy and timeliness of desertification risk prediction. Future research could further explore model optimization and multi-dimensional data fusion to address dynamic changes in complex environments, promoting the implementation of precise ecological management strategies.

References

1. Imbrenda V ,Maialetti M ,Sateriano A , et al.Scrutinizing the Statistical Distribution of a Composite Index of Soil Degradation as a Measure of Early Desertification Risk in Advanced Economies[J].Environments,2024,11(11):246-246.
2. Zotos A ,Kokkoris P I ,Charalampopoulos I , et al.Wetlands in Crisis: The Silent Desertification Threat on the Greek Wetlands[J].Land,2024,13(10):1567-1567.
3. Bruno E ,Salvia R ,Quaranta G , et al.Identifying On-Site and Off-Site Drivers of Land Degradation in Advanced Economies: A Spatial Approach for Italy[J].Environmental and Resource Economics,2024,87(9):2429-2453.
4. Badiee H S ,Masoudi M ,Abtahi A S , et al.Assessment of desertification hazard using RS and GIS to compare the efficiency of FAO-UNEP and MEDALUS models[J].Spatial Information Research,2024,32(5):641-650.
5. Hamad O K ,Surucu A .Land degradation sensitivity and desertification risk in Harrir region, northern Iraq[J].Heliyon,2024,10(5):e27123-.
6. Honghui Z ,Meiling D ,Xiaojun Y .Oasification in Arid and Semi-Arid Regions of China: New Changes and Re-Examination[J].Sustainability,2023,15(4):3335-3335.
7. Luca S .Framing socioecological complexity: The long-term evolution of multiple dimensions of desertification risk in Italy.[J].Risk analysis : an official publication of the Society for Risk Analysis,2022,43(8):1657-1666.

8. Vito I ,Rosa C ,Valerio S D , et al.Modeling Spatio-Temporal Divergence in Land Vulnerability to Desertification with Local Regressions[J]. Sustainability, 2022, 14(17): 10906-10906.
9. Ali M ,Rabeh G ,Daoud W .Mapping potential desertification-prone areas in North-Eastern Algeria using logistic regression model, GIS, and remote sensing techniques[J]. Environmental Earth Sciences,2022,81(15):385-385.
10. Margherita C ,Rosanna S ,Giovanni Q , et al.Official statistics, spatio-temporal dynamics and local-scale monitoring: toward integrated environmental-economic accounting for land degradation[J].Letters in Spatial and Resource Sciences,2022,15(3):469-491.