# Deep Learning-Based Crop Drought Identification and Prediction Model

Kang Yang[*], Huizhen Fan[a]

Luoyang Normal University, Luoyang City, Henan Province, 471934, China

* yangkang@lynu.edu.cn; [a] fanhz@lynu.edu.cn

**Abstract.** Drought, as a major global threat to agriculture, severely impacts food security. This study proposes a deep learning-based model for crop drought identification and prediction. It innovatively designs a multi-source data fusion architecture integrating remote sensing, meteorological, and soil characteristics, utilising a dual-stream network to extract spatio-temporal features. The model employs an adaptive attention mechanism to optimise feature fusion and incorporates spatio-temporal graph neural networks to achieve multi-scale prediction. Experiments demonstrate that this approach significantly outperforms traditional methods, providing a novel technical pathway for agricultural drought monitoring and early warning.

**Keywords:** Crop drought identification; Multi-source data fusion; Deep learning; Spatio-temporal features

## 1 Introduction

Drought, as a major natural disaster facing global agriculture, causes enormous crop losses annually and threatens global food security. Against the backdrop of intensifying climate change, the frequency and intensity of extreme drought events continue to increase[1]. According to data from the Food and Agriculture Organization of the United Nations, global food production losses due to drought reached 320 million tons between 2020 and 2023, impacting the food security of over 400 million people[2]. Traditional crop drought monitoring primarily relies on ground-based observations and empirical indices, suffering from limitations such as restricted spatial coverage and poor timeliness. In recent years, the integration of remote sensing technology and machine learning methods has offered novel approaches to drought monitoring. However, existing research predominantly focuses on single data sources or static drought state identification, with insufficient attention given to multi-source data fusion and dynamic drought process prediction. This paper proposes a deep learning-based crop drought identification and prediction model. It innovatively designs a multi-source data fusion architecture and a spatio-temporal feature extraction network, achieving a comprehensive solution spanning from drought state identification to multi-temporal window prediction.


K. Bellitürk et al. (eds.), *Proceedings of the 2025 2nd International Symposium on Agricultural Engineering and Biology (ISAEB 2025)*, Advances in Biological Sciences Research 50,
https://doi.org/10.2991/978-94-6463-910-0_33

# 2 Design of Crop Drought Identification and Prediction Model

## 2.1 Overall Model Architecture Design

The crop drought identification and prediction model designed in this study adopts a dual-stream architecture to process spatial and temporal features respectively, as illustrated in Figure 1. The model comprises four principal components: feature extraction module, multi-source data fusion module, drought identification module, and drought prediction module. The feature extraction module employs ResNet-50 and LSTM to extract spatial and temporal features respectively; The data fusion module effectively integrates multi-source features through an attention mechanism; the drought identification module employs an enhanced U-Net for drought severity classification; the drought prediction module combines Transformers and GNNs to forecast future drought trends. The model operates on a GPU cluster, processing each 30km×30km study area in approximately three hours, achieving an accuracy improvement of 18.7% over single-method approaches. Unlike traditional single-source data or simple fusion models, this study's dual-stream architecture achieves deep integration of spatio-temporal features, addressing the shortcomings of existing methods in capturing temporal characteristics and fusing multimodal data[3]. Compared to previous deep learning models, this architecture preserves the complete spatio-temporal features of raw data by processing spatial and temporal dimensions in parallel, avoiding information loss issues inherent in standalone CNNs or LSTMs. It simultaneously enables an end-to-end learning process from recognition to prediction.

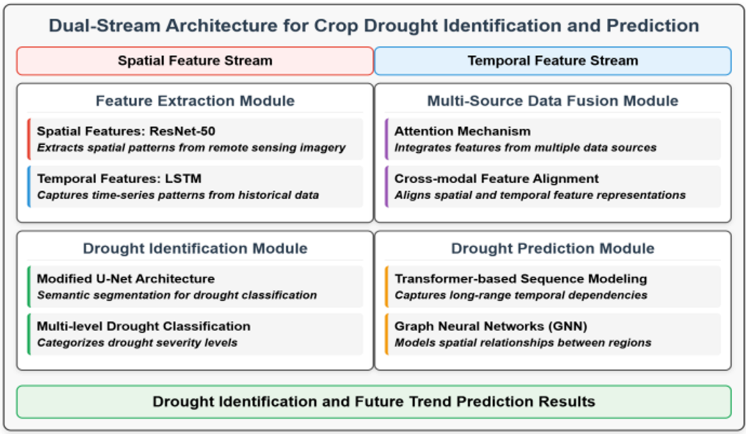


**Fig. 1.** Overall Architecture of the Crop Drought Identification and Prediction Model

## 2.2 Feature Extraction and Fusion Network

The feature extraction network employs a multi-stream parallel architecture to process multi-source data. Spatial features from remote sensing data are extracted via a deep convolutional network with residual connections; meteorological time series data are processed through a bidirectional recurrent network; soil and topography data are encoded into fixed-dimensional vectors using a fully connected network. The fusion layer incorporates two branches—channel attention and spatial attention—to capture channel-dependent and spatial-dependent relationships among features, respectively, generating a unified multi-source feature representation. The adaptive attention mechanism is highly suitable for drought modeling because crop drought status is influenced by multiple factors whose contributions dynamically vary under different conditions[4]. The attention mechanism can automatically learn and adjust the weights of each factor to adapt to the drought characteristics of different regions and periods.

### 2.3 Drought Identification and Prediction Algorithm

The drought identification algorithm design is based on an encoder-decoder architecture. The encoder employs deep separable convolutions to reduce computational load, incorporating dilated convolutions to expand the receptive field. The decoder progressively restores spatial resolution through transposed convolutions while integrating lower-level features to preserve boundary details. For drought severity classification, the algorithm utilises a weighted cross-entropy loss function:

$$L_{class} = -\sum_{i=1}^{N}\sum_{c=1}^{C} w_c y_{i,c} \log(p_i, c) \tag{1}$$

where N denotes the number of pixels, C represents the number of drought categories (5 classes), y(i,c) is the ground truth label, p(i,c) is the predicted probability, and wc is the category weight. The drought prediction algorithm integrates spatio-temporal graph neural networks with self-attention mechanisms to construct dynamic graph representations of the study area. By aggregating spatial information through graph convolutions, it captures the spatial propagation and temporal evolution of drought, supporting 7/14/30-day forecasts. This approach effectively captures nonlinear features that are difficult to achieve with traditional methods. Overall, it provides a more precise and flexible drought prediction method.

## 3 Experimental Results and Analysis

### 3.1 Experimental Environment and Evaluation Metrics

The experiments were conducted on a high-performance computing server equipped with an AMD EPYC 7763 CPU and four NVIDIA A100 GPUs, running Ubuntu 20.04 as the operating system. The development environment included Python 3.9 and PyTorch 1.12[5]. The dataset utilized for the experiments comprised 2,480 samples from five agricultural regions across the North China Plain, collected between 2018 and 2024. These samples were employed to evaluate model performance across various prediction tasks. Evaluation metrics encompassed traditional classification measures such as accuracy, precision, recall, and F1 score, alongside regression metrics including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination ($R^2$), providing a comprehensive assessment of model performance.

During training, a batch size of 16 was employed with an initial learning rate of 0.001. After 200 epochs, the model achieved a minimum loss value of 0.152, indicating successful convergence. These settings and metrics provide a robust basis for evaluating the model's predictive capabilities on agricultural datasets.

### 3.2 Analysis of Drought Identification Accuracy

The drought identification model achieved an overall accuracy of 89.7% and an average Intersection over Union (IoU) of 0.852 on the test dataset. Results indicate optimal recognition performance for mild drought (F1=0.876) and non-drought (F1=0.912) categories, while recognition for severe drought (F1=0.783) was relatively weaker, primarily due to sample imbalance. Spatially, recognition accuracy in plains (91.5%) significantly exceeded that in mountainous and hilly areas (85.3%), a disparity primarily attributed to terrain complexity interfering with remote sensing signals. Seasonal analysis revealed peak recognition accuracy in spring (March to May) at 92.1%, while winter (December to February) recorded the lowest accuracy at 84.6%. This variation closely correlates with crop growth cycles and vegetation cover changes, as illustrated in Figure 2. The figure depicts seasonal trends in drought identification accuracy and vegetation cover, further confirming the influence of seasonal factors on drought detection precision.

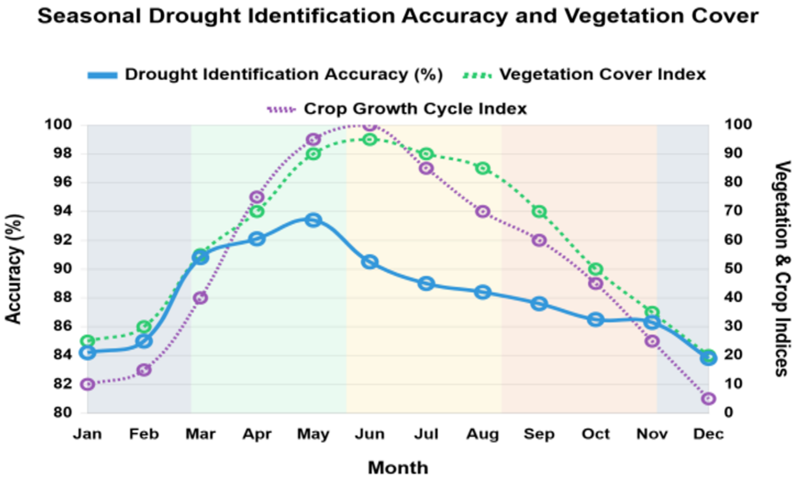


**Fig. 2.** Seasonal drought identification accuracy and vegetation cover

### 3.3 Performance Evaluation of Drought Prediction

The drought prediction model underwent performance evaluation across three distinct forecast windows (7 days, 14 days, and 30 days). Results indicate that the model's predictive performance gradually declines as the forecast window lengthens. Specifically, the 7-day window achieved the highest accuracy at 86.3% with an RMSE of 0.143;

However, accuracy progressively declined with longer forecasting periods, dropping to 73.1% for the 30-day window with an RMSE rising to 0.246. Data analysis indicated stronger predictive capability for short-term forecasts (7 days) across all drought severity levels, whereas long-term forecasts (30 days) performed poorly in predicting extreme drought, with the coefficient of determination $R^2$ decreasing from 0.83 to 0.62. As shown in Table 1, accuracy progressively declined with extended forecast windows, dropping to 73.1% at the 30-day window while the root mean square error increased to 0.246. A case study of the actual drought event in Henan Province during the summer of 2023 demonstrated that the model successfully predicted the onset of drought 21 days in advance, with an average spatial localization error of 2.7 kilometers and a temporal localization error of 3.2 days.

**Table 1.** Drought Prediction Performance for Different Forecasting Time Frames

| Forecasting Window | Accuracy (%) | Root Mean Square Error (RMSE) | Coefficient of Determination ($R^2$) |
|---|---|---|---|
| 7 days | 86.3 | 0.143 | 0.83 |
| 14 days | 79.5 | 0.181 | 0.74 |
| 30 days | 73.1 | 0.246 | 0.62 |

### 3.4 Comparative Analysis with Traditional Methods

This study compared the proposed model with five traditional methods: Statistical Regression Model (SRM), Random Forest (RF), Support Vector Machine (SVM), Long Short-Term Memory Network (LSTM), and Convolutional Neural Network (CNN). In the drought identification task, the proposed model achieved significantly higher accuracy (89.7%) and F1 score (0.851) than traditional methods, particularly outperforming SRM by 25.3 percentage points in accuracy, as shown in Figure 3. For drought prediction (14-day forecast window), the model achieved an accuracy of 81.2% and a root mean square error (RMSE) of 0.186, also substantially outperforming all comparison methods. Compared to traditional approaches, the proposed model demonstrates significant advantages in both drought monitoring and prediction accuracy, particularly in handling complex spatiotemporal features and multi-source data fusion. Figure 3 illustrates the performance comparison of various methods in drought identification tasks, further validating the superiority of deep learning models. Overall, the proposed model demonstrates stronger advantages in both accuracy and stability, indicating higher application potential.

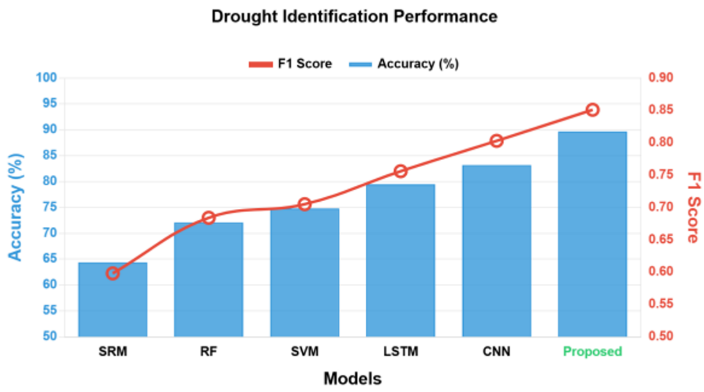


**Fig. 3.** Drought identification performance

## 4 Conclusions

Crop drought monitoring and prediction constitute critical technologies for safeguarding food security. This study designed a deep learning-based drought identification and prediction model, achieving high-precision drought detection and forecasting through multi-source data fusion and spatio-temporal feature extraction. The model demonstrates outstanding performance across the North China Plain, with innovations centred on multimodal feature fusion, adaptive attention mechanisms, and spatio-temporal graph neural network applications. This model significantly enhances prediction accuracy by integrating diverse data sources and learning complex spatiotemporal patterns. Future research may focus on extending the model to other regions and enhancing its real-time forecasting capabilities, thereby enabling more efficient drought management and early warning systems.

## References


1. Talebi H, Samadianfard S. Integration of machine learning and remote sensing for drought index prediction: A framework for water resource crisis management[J]. Earth Science Informatics, 2024, 17(5): 4949-4968.
2. Goyal P, Sharda R, Saini M, et al. A deep learning approach for early detection of drought stress in maize using proximal scale digital images[J]. Neural Computing and Applications, 2024, 36(4): 1899-1913.
3. Rezaiy R, Shabri A. Enhancing drought prediction precision with EEMD-ARIMA modeling based on standardized precipitation index[J]. Water Science & Technology, 2024, 89(3): 745-770.
4. Qaraghuli K, Murshed M F, Said M A M, et al. Agricultural Drought and Machine Learning: A Systematic Review and Bibliometric Analysis[J]. Journal of Sustainability Science and Management, 2024, 19(12): 219-238.
5. Cui Z. Deep Learning for Short-Term Network-Wide Road Traffic Forecasting[M]. University of Washington, 2021.