Article

Utilizing Deep Learning and Spatial Analysis for Accurate Forest Fire Occurrence Forecasting in the Central Region of China

Youbao Guo 1,2, Quansheng Hai 1,3,4,* and Sainbuyan Bayarsaikhan 1,4,5

1 Department of Geography, School of Arts and Sciences, National University of Mongolia, Ulaanbaatar 14200, Mongolia
2 Department of Railway Engineering, Baotou Railway Vocational and Technical College, Baotou 014060, China
3 Department of Ecology and Environment, Baotou Teacher’s College, Baotou 014030, China
4 Laboratory of Geoinformatics (GEO-iLAB), Graduate School, National University of Mongolia, Ulaanbaatar 14200, Mongolia
5 Research Institute of Urban and Regional Development, National University of Mongolia, Ulaanbaatar 14200, Mongolia
* Correspondence: 66063@bttc.edu.cn

Abstract: Forest fires in central China pose significant threats to ecosystem health, public safety, and economic stability. This study employs advanced Geographic Information System (GIS) technology and Convolutional Neural Network (CNN) models to comprehensively analyze the factors driving the occurrence of these fire events. A predictive model for forest fire occurrences has been developed, complemented by targeted zoning management strategies. The key findings are as follows: (i) Spatial analysis reveals substantial clustering and spatial autocorrelation of fire points, indicating high-density areas of forest fire occurrence, primarily in Hunan and Jiangxi provinces, as well as the northeastern region. This underscores the need for tailored fire prevention and management approaches. (ii) The forest fire prediction model for the central region demonstrates exceptional accuracy, reliability, and predictive power. It achieves outstanding performance metrics in both training and validation sets, with an accuracy of 86.00%, precision of 88.00%, recall of 87.00%, F1 score of 87.50%, and an AUC value of 90.50%. (iii) Throughout the year, the occurrence of forest fires in central China varies by location and season. Low-occurrence periods are observed in summer and winter, particularly in Hunan and Hubei provinces, due to moderate weather conditions, agricultural practices, and reduced outdoor activities. However, spring and autumn also present localized risks due to uneven rainfall and dry climates. This study provides valuable insights into the dynamics of forest fire occurrences in central China, offering a solid framework for proactive fire management and policy formulation to effectively mitigate the impacts of these events.

Keywords: forest fires; central China; autocorrelation; convolutional neural network (CNN); geographic information system (GIS); prediction accuracy

1. Introduction

Forests, Earth’s verdant repositories, are essential for sustaining ecological equilibrium, fostering biodiversity, regulating climate change, and bolstering human welfare. They serve as significant carbon sinks, purifying the air, safeguarding water sources, and providing habitats for myriad species. Additionally, forests offer humans crucial resources such as food, medicine, and livelihood opportunities [1–4]. Forest fires exhibit a dual nature in natural ecosystems. On the positive side, moderate fires contribute to ecological balance by fostering germination conditions for specific plant seeds, eliminating diseased trees, and enriching biodiversity [5–10]. However, large-scale forest fires can wreak havoc on ecological environments, diminishing carbon sinks, imperiling wildlife habitats, causing biodiversity loss, and triggering severe air pollution and accelerated climate change. Accurate forecasting of forest fires is paramount. It enables the proactive deployment of preventive strategies.
measures, minimizing casualties and property damage, while safeguarding the ecological environment and biodiversity, thereby mitigating the adverse impacts of fires on climate change [11,12]. Through scientific prediction and timely response, resource allocation can be optimized, and emergency response efficiency augmented, playing a pivotal role in upholding public safety, conserving natural resources, and fostering sustainable economic and social development [13–17].

Forest fire prediction is a complex challenge influenced by dynamic factors, such as climate, terrain, vegetation, and human activities, requiring diverse and high-quality data sources for accurate predictions. Various models, including physical, statistical learning, and machine learning approaches, each offer distinct advantages in forest fire prevention and control [18]. Physical models delve into the fundamental mechanisms of fire spread, encompassing heat conduction, convection, and radiation. By simulating combustion dynamics within forests and incorporating variables such as weather, terrain, and vegetation types, these models are able to forecast fire behavior and evolution [19–25]. However, despite their in-depth insights, physical models often require significant computational resources and detailed environmental data, limiting their practicality for real-time or large-scale predictions [26–29]. On the other hand, statistical learning models leverage historical fire data and the relationships among environmental factors to predict future fire occurrences [30–34]. Techniques such as regression and time-series analysis are employed to discern patterns and trends [35]. However, due to the complexity of forest fires and the sheer diversity of influencing factors, purely statistical models sometimes fall short in capturing all relevant dynamic changes.

In contrast, machine learning models, including Random Forests, Support Vector Machines (SVMs), GBDT (Gradient Boosting Decision Tree), MLP (Multi-Layer Perceptron), and so on [9,36–41], and particularly deep learning models, have emerged as powerful tools for forest fire prediction. Deep learning, especially Convolutional Neural Networks (CNNs), is widely used in image recognition and natural language processing due to its ability to handle large datasets, recognize complex patterns, and automatically extract features [42–44]. CNNs can automatically identify and learn the intricate nonlinear relationships between fire influencing factors without manual intervention, extracting key features from satellite images, meteorological data, and topographical information [13,45,46]. This capability not only enhances prediction accuracy but also allows for the processing of vast amounts of data, enabling rapid, real-time fire prediction. Additionally, the application of CNN models improves the generalizability of prediction models, making them adaptable to fire prediction needs in various regions and conditions [13,47,48].

Preventing the occurrence of forest fires is critically important, especially in the northeastern and central provinces of China’s central region. These areas, characterized by extensive forest cover, serve as crucial biodiversity reservoirs, essential for water conservation, and act as vital ecological safety barriers. Forest fires can cause significant ecological damage, destroy wildlife habitats, and endanger the lives and property of local communities. In severe cases, they can also disrupt regional climate patterns and impact the global carbon cycle [49]. Therefore, enhancing forest fire occurrence prevention in these areas is crucial for safeguarding the ecological environment, preserving biodiversity, ensuring socio-economic stability, and promoting regional climate balance.

The objectives of this study are as follows: (i) Analyzing the patterns of forest fire occurrences over a 20-year period using Geographic Information System (GIS) techniques. (ii) Predicting forest fire occurrences using a Convolutional Neural Network (CNN) model that integrates meteorological, topographical, socio-economic, and vegetation factors. (iii) Forecasting and mapping seasonal forest fire occurrences to develop tailored prevention strategies.

2. Resources and Methods
2.1. Study Area

The central region of China spans from the northeast, through the north and central parts, to the southeastern area, encompassing the provinces of Heilongjiang, Jilin, Shanxi,
Henan, Anhui, Hubei, Jiangxi, and Hunan (Figure 1). This region showcases diverse climates, ranging from temperate to subtropical, and features varied terrains including plains, hills, and mountains. It boasts some of China’s most fertile black soil and diverse soil types, supporting extensive agricultural production. Major rivers such as the Yellow River and the Yangtze River traverse this area, providing ample water resources, making it a crucial grain production base and a hub for hydroelectric power [50,51].

The region’s extensive forest cover, from the coniferous forests in the north to the broadleaf mixed forests in the south, serves as a reservoir of biodiversity and a source of timber and non-timber forest products. Socio-economically, the central region is characterized by a balance of agriculture and industry, hosting both old industrial bases focused on heavy industry and major agricultural provinces. In recent years, there has been significant development in tourism and high-tech industries, contributing to China’s economic growth. However, the complex natural conditions of this region also make it susceptible to forest fires. Due to its unique geographical location and natural conditions, the central region is a key area for studying various issues related to China’s natural environment, socio-economic development, and ecological conservation [52].

2.2. Data Sources

In this study, by utilizing the MODIS (Moderate Resolution Imaging Spectroradiometer) data provided by the National Aeronautics and Space Administration (NASA)
(https://earthdata.nasa.gov/firms, accessed on 12 October 2023) [53–55], fire occurrence records from January 2001 to December 2019 were analyzed, focusing on vegetation fire points with a confidence level exceeding 80%. These data integrate detections from both the Aqua and Terra satellites and are based on the fire location information generated by the MOD14/MYD14 products. The information includes detailed data such as ignition time, location coordinates, fire type, and confidence level, with a spatial resolution of 1 km. This allows for the identification of fires and other thermal anomalies [56]. VIIRS (Visible Infrared Imaging Radiometer Suite) is a satellite sensor jointly developed by NASA and NOAA, designed to capture visible and infrared imagery of the Earth’s surface globally. With its high spatial resolution of 375 m and multispectral observation capabilities, VIIRS plays a crucial role in monitoring forest fires. Using infrared and thermal infrared data, VIIRS can accurately detect and track surface fire points, providing essential support for fire alert systems and emergency decision making [57,58].

This study adopted ArcGIS 10.2. Firstly, a point layer of fire spots was created, and then duplicate fire spots less than 1 km apart within 24 h were eliminated, with only the initial fire spots retained. These fire spots were then overlaid with China’s administrative division vector map and forest cover map to exclude fire spots located in non-forest areas and outside China’s administrative regions, thus accurately identifying forest fire spot information in various provinces and cities in China to support forest fire monitoring and management efforts. Therefore, a total of 25,412 fire points were detected within the forest cover area from MODIS data spanning 2001 to 2019, while VIIRS data from 2012 to 2019 identified 5874 fire points within the same forest cover area.

The vegetation type data were sourced from the Resource and Environment Data Center of the Chinese Academy of Sciences, available at https://www.resdc.cn/ (accessed on 1 October 2023) [59]. This dataset primarily utilizes Landsat satellite imagery from the United States to construct a national-scale multi-temporal land use and land cover thematic database for China through manual visual interpretation. The interpretation process integrates various image features such as texture, tone, and shape to establish interpretation markers, employing an object-oriented approach combined with manual visual interpretation to derive the final results [60]. Regarding forest land classification, the dataset includes four main types, achieving an average classification accuracy of over 80%: forested areas, shrublands, sparse forests, and other forested lands.

Based on a comprehensive literature review and extensive field investigation [18,61,62], this study analyzes various factors crucial for understanding the ecosystem. Despite lightning being recognized as an important potential trigger, acquiring accurate and comprehensive lightning data through specialized equipment and complex data processing procedures exceeded our budgetary and temporal constraints. Therefore, we focused our analysis on other more feasible and crucial factors, such as climate patterns (e.g., temperature, humidity, precipitation), topographical features (e.g., slope, elevation, soil type), vegetation composition (e.g., tree species distribution), and socio-economic conditions (e.g., population density, road networks, fire prevention policy implementation). To validate and enrich the findings from the literature, we performed a rigorous field investigation. Through site visits and interviews, we gathered firsthand information on the various factors within the specific study area. Our observations highlighted significant variations in climatic conditions, topographical characteristics, and vegetation distributions, all of which impacted the frequency and severity of forest fires. Furthermore, we found that socio-economic conditions played a crucial role in shaping fire prevention and response capabilities, with densely populated areas often demonstrating stronger fire awareness and more effective emergency response mechanisms.

In summary, our study, grounded in both the literature review and field investigation, delved into the key factors influencing forest fire occurrences and aimed to provide practical and targeted recommendations for relevant decision makers to mitigate the impacts of forest fires. These factors, including climate patterns, topographical features, vegetation
composition, and socio-economic conditions, are meticulously detailed in Table 1 below. The main relevant figures are shown in Figure 2 and Appendix A.

Table 1. Summary of the data sources utilized in the study.

<table>
<thead>
<tr>
<th>Classification</th>
<th>Data</th>
<th>Resolution</th>
<th>Source</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topographic</td>
<td>Slope/Elevation/Slope direction</td>
<td>1 km</td>
<td><a href="https://www.resdc.cn">https://www.resdc.cn</a> (accessed on 10 August 2023)</td>
<td>[63,64]</td>
</tr>
<tr>
<td>Climate</td>
<td>Average daily surface temperature/average daily relative humidity/daily maximum surface temperature, etc.</td>
<td>-</td>
<td><a href="https://data.cma.cn">https://data.cma.cn</a> (accessed on 8 August 2023)</td>
<td>[12,65,66]</td>
</tr>
<tr>
<td>Vegetation</td>
<td>Vegetation types</td>
<td>-</td>
<td><a href="https://www.resdc.cn">https://www.resdc.cn</a> (accessed on 9 September 2022)</td>
<td>[67,68]</td>
</tr>
<tr>
<td>Social and human factors</td>
<td>Distance from road/Distance from residential area/Gross Domestic Product/Population</td>
<td>1:100,000, 1:100,000, 1 km, 1 km</td>
<td><a href="https://www.resdc.cn">https://www.resdc.cn</a> (accessed on 12 September 2023)</td>
<td>[69,70]</td>
</tr>
</tbody>
</table>

Figure 2. Cont.
Figure 2. The main data chart used in this article.

2.3. Method

In this comprehensive study, Figure 3 plays a pivotal role by detailing the technical pathway used to explore the complexities surrounding forest fires. This roadmap meticulously explains the integration process of various datasets, offering multiple perspectives on forest fire phenomena. These datasets encompass a wide array of information, including detailed fire records, land use patterns, meteorological observations, socio-economic factors, vegetation descriptions, and topographical data.

To ensure the comparability and analyzability of these diverse data sources, advanced normalization techniques are employed. These techniques effectively reduce amplitude variations among datasets, creating a unified data framework that supports consistent and balanced analysis. During the data preparation stage, we conducted detailed analyses employing advanced data inspection techniques. We utilized kernel density analysis to identify areas with a high concentration of fire incidents, effectively pinpointing forest fire hotspots. Furthermore, we performed spatial autocorrelation analysis to uncover complex spatial relationships between fire events, gaining valuable insights into their interconnectedness.

To further deepen these analytical insights, this study employs an advanced Convolutional Neural Network (CNN) deep learning algorithm. This machine learning approach integrates various variables, such as historical fire data, meteorological conditions, land use, and socio-economic factors, to accurately predict forest fire. The predictive model provides decision makers with essential information to develop targeted forest fire prevention and control strategies.

Thus, in the Methods section, Figure 3 is explained to highlight the main steps presented in the roadmap, including the integration of datasets, application of normalization techniques, detailed data analysis methods (such as kernel density and spatial autocorrelation analyses), and the use of CNN deep learning for accurate forest fire prediction.
2.3.1. Kernel Density Estimation (KDE)

Kernel density analysis is a spatial statistical technique that estimates the density distribution of data points in space by placing a kernel function around each data point and summing the contributions [71–73].

In the context of forest fire prediction, this method utilizes a large dataset of historical fire points to visually represent the spatial distribution density of fires using smoothing.
kernel functions. It identifies high-risk areas prone to fires without assuming the data’s distribution form beforehand, capturing complex interactions among factors. This non-parametric approach provides a more accurate basis for forest fire prediction. Bandwidth is a crucial parameter in kernel density estimation (KDE) as it directly influences the smoothness and accuracy of the density estimate. In KDE, the bandwidth determines the contribution of each data point to the final density estimate, thereby affecting the peaks and shape of the estimate. The choice of bandwidth is typically a trade-off. A bandwidth that is too small can lead to overfitting the data, resulting in excessive fluctuations and noise, which hinders capturing the true characteristics of the data distribution—this is known as overfitting. Furthermore, its robust spatial analysis capabilities help reveal the dynamic changes of forest fires, supporting the development of effective prevention and control strategies.

The announcement is as follows [74].

\[ f(x) = \sum_{i=1}^{n} \frac{1}{\pi r^2} \Phi \left( \frac{d_{ix}}{r} \right), \] (1)

In the equation, “r” represents the search radius, “n” denotes the total number of forest fire sample points, “d_{ix}” signifies the distance between forest fire point “i” and “x”, and “Φ” denotes the distance weight.

2.3.2. Spatial Autocorrelation Analysis

Spatial autocorrelation is an important concept in Geographic Information Systems (GISs) and spatial statistics, used to analyze patterns and trends in spatial data [75–77]. It measures the degree of correlation between values at different points in geographic space. Global spatial autocorrelation and local spatial autocorrelation are two common methods of spatial autocorrelation. Spatial autocorrelation plays a crucial role in wildfire research. Through global spatial autocorrelation, we can identify the overall distribution pattern of wildfires, understand whether they exhibit clustering tendencies, and determine the location and scale of clusters. Local spatial autocorrelation, on the other hand, reveals hotspot areas of wildfires, aiding decision makers in more accurately allocating resources and formulating prevention and control strategies. These analytical findings help us gain a deeper understanding of the spatial heterogeneity of wildfires and provide scientific evidence and guidance for wildfire management and prevention.

The formulas are as follows [78,79]:

Global autocorrelation:

\[ I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} (x_i - \bar{x})(x_j - \bar{x})}{n \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} (x_i - \bar{x})^2}, \] (2)

In this equation, \( I \) signifies the global Moran’s I index, \( n \) indicates the total number of spatial units, \( W_{ij} \) represents the spatial weights between units \( i \) and \( j \), \( x_i \) and \( x_j \) denote the values of variable \( x \) for units \( i \) and \( j \), and \( \bar{x} \) denotes the mean or average of variable \( x \).

Local autocorrelation:

\[ I' = \frac{[n(x_i - \bar{x}) \sum_{j=1}^{n} W_{ij} (x_j - \bar{x})] / \sum_{i=1}^{n} (x_i - \bar{x})^2}{(x_i - \bar{x})^2}, \] (3)

In this formula, \( I' \) denotes the local Moran’s I index, \( n \) refers to the count of spatial units, \( W_{ij} \) indicates the spatial weights between units \( i \) and \( j \), \( x_i \) is the value of variable \( x \) for unit \( i \), and \( \bar{x} \) signifies the mean of variable \( x \).

In this equation, \( I' \) stands for the local Moran’s I index, \( n \) represents the total number of spatial units, \( W_{ij} \) designates the spatial weights between units \( i \) and \( j \), \( x_i \) denotes the value of variable \( x \) for unit \( i \), and \( \bar{x} \) indicates the average of variable \( x \).
2.3.3. Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) draw inspiration from the human visual system, particularly how our brains process and interpret optical signals. In the human visual system, different neurons have specific responses to different features in images, such as edges, angles, and colors. Similarly, CNNs simulate this hierarchical processing mechanism through their multi-layered structure, enabling them to automatically learn and extract features from simple to complex in images [80, 81]. CNNs demonstrate powerful capabilities in extracting key information, reducing data dimensions, and recognizing patterns in complex backgrounds. In the application of predicting forest fire occurrences, CNNs can automatically learn and identify key features from satellite images and other geospatial data that may indicate fire risks. They can identify complex, fire-related patterns and features, thereby improving prediction accuracy [82, 83].

The ability to process large amounts of image data and quickly identify potential fire risk areas is particularly useful for real-time or near-real-time monitoring. As more data are collected, CNN models can continue to learn and adapt to identify new fire patterns [13].

The convolutional layer applies a set of filters (kernels) to the input data to extract spatial patterns. Let us denote: input image as $X$ with dimensions $W \times H \times D$, where $W$ is width, $H$ is height, and $D$ is the number of channels (e.g., RGB channels for color images).

A filter (kernel) as $K$ with dimensions $F \times F \times D$, where $F$ is the filter size.

For a convolution operation at a specific spatial location $(i, j)$ in the input:

$$ (X * K)_{ij} = \sum_{m=0}^{F-1} \sum_{n=0}^{F-1} \sum_{d=0}^{D-1} X_{(i+m)(j+n)d}K_{mn} $$

This operation is applied across the entire image, producing an output feature map.

Batch Normalization (BN) normalizes the activations of a convolutional layer to stabilize and accelerate training. It computes:

$$ \hat{X} = \frac{x - \mu}{\sqrt{\sigma^2 + \epsilon}}, $$

where $\mu$ and $\sigma^2$ are the mean and variance over the mini-batch, and $\epsilon$ is a small constant to avoid division by zero.

The LeakyReLU activation function introduces a small slope for negative input values to prevent dying ReLU problem:

$$ \text{LeakyReLU}(X) = \begin{cases} 
X, & \text{if } X \geq 0 \\
\alpha X, & \text{if } X < 0 
\end{cases} $$

where $\alpha$ is a small constant.

As depicted in Figure 4, a meticulously designed 8-layer Convolutional Neural Network (CNN) was created for predicting forest fire occurrences. The network consists of Convolutional Layers (Conv layers), Batch Normalization (BN layers), and utilizes the LeakyReLU activation function. To train and validate the model, the dataset was partitioned into a 7:3 ratio. Extensive parameter configuration was performed, setting the learning rate at 0.001, weight decay at 0.01, momentum at 0.9, and a fixed L1 regularization factor of 0.01 [18]. Parameter optimization was carried out using the stochastic gradient descent (SGD) optimization algorithm. These parameter settings and the CNN architecture were carefully crafted to optimize the model’s performance in predicting forest fire occurrences. For further details, refer to Appendix B.
3.1. The Results of Kernel Density Estimations

Figure 5 illustrates that the central region exhibits a high density of forest fire occurrence, as indicated by kernel density analysis. These high-risk areas are predominantly concentrated in Hunan Province, particularly in Shaoyang, Hengyang, and Yongzhou. Additionally, Jiangxi Province, including Ji’an and Ganzhou, also shows a significant density of forest fire events. Similarly, the northeastern regions of Greater Khingan, Heihe, and Yichun are identified as high-fire-occurrence areas. The elevated fire occurrence in these regions may be attributed to unique climatic conditions, vegetation types, and human activities.

In Hunan Province, Shaoyang, Hengyang, and Yongzhou face increased probabilities of fire occurrence due to their abundant forest resources and relatively high population density. The local climatic conditions, characterized by high temperatures and dry seasons, further exacerbate the occurrence of fire. Similarly, Ji’an and Ganzhou in Jiangxi Province exhibit complex terrains and diverse vegetation types, coupled with frequent human activities such as agricultural production and tourism development, contributing to heightened forest fire occurrence.

In the northeastern regions of Greater Khingan, Heihe, and Yichun, despite lower population density, extensive forest resources and cold, dry climatic conditions render these areas prone to high occurrences of forest fires. Particularly during spring and autumn, dry and windy conditions make fires challenging to contain once ignited.

2.3.4. Evaluation Indicators

In machine learning and statistical analysis, a set of key metrics—accuracy, precision, recall, F1 score, and AUC (Area Under the Curve)—are essential for assessing the performance of classification models. The formulas for these metrics are as follows [18,84–86]:

\[ \text{Recall} = \frac{TP}{TP + FN}, \]  
\[ \text{Precision} = \frac{TP}{TP + FP}, \]  
\[ \text{Accuracy} = \frac{TP + TN}{(TP + FP + TN + FN)}, \]  
\[ F1 = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \]

*Accuracy* measures the overall correctness of predictions but can be misleading with imbalanced datasets. *Precision* highlights the proportion of correctly predicted fires out of all fire predictions, crucial when false alarms are costly. *Recall* focuses on detecting actual fires, minimizing missed detections. The F1 score balances precision and recall, providing a comprehensive performance metric. Lastly, AUC evaluates the model’s ability to distinguish between fire and non-fire instances across various thresholds, offering a robust measure of overall performance. Each metric has its strengths and weaknesses, and their combined use ensures a thorough assessment of the prediction model’s effectiveness.
3.2. Results of Autocorrelation Analysis on Forest Fire Occurrences in Central China Region

Figure 6 illustrates the spatial distribution of forest fire occurrences and their global and local spatial clustering characteristics in the central region of China. A total of 19 cities show a high–high (H_H) global autocorrelation, indicating significant spatial clustering and a relatively high incidence of forest fires. These cities are concentrated primarily in Hunan Province, specifically in Shaoyang, Hengyang, and Yongzhou, as well as in Jiangxi Province (Ji’an and Ganzhou) and the northeastern region (Daxing’anling, Heihe, and Yichun).

This clustering of forest fires can be attributed to several factors. Firstly, the region’s climate, characterized by long dry seasons and limited precipitation, creates conditions favorable for fire ignition and spread. Secondly, the presence of highly flammable vegetation types such as pine and broadleaf forests further exacerbates the risk. Additionally, human activities such as agricultural burning and deforestation contribute significantly to the occurrence of fires.

Furthermore, three cities exhibit a high–low (H_L) local autocorrelation pattern, located in Hubei Province, such as Suizhou and Huanggang. This suggests that while these cities have a high forest fire risk individually, their correlation with neighboring cities is relatively weak. This variation may stem from differences in topography, land use patterns, or local fire management strategies.

Additionally, 14 cities display a low–high (L-H) pattern, distributed across Hunan Province (Yiyang, Loudi), Jiangxi Province (Shangrao), and Yichun in the northeast. This indicates that while these cities themselves may face a relatively low forest fire risk, their surrounding areas exhibit a higher risk. Addressing this situation requires enhanced cooperation between these cities and their neighboring regions to collectively prevent and manage forest fires.
Notably, the cities exhibiting local H_H autocorrelation are clustered predominantly in Hunan Province (Hengyang, Shaoyang), Ganzhou in Jiangxi, and the northeastern region (Daxing’anling, Heihe, Yichun). Only three cities show an L-H pattern, while the majority are classified as low–low (L_L) or areas with insignificant autocorrelation.

In summary, Figure 6 provides valuable insights into the spatial distribution of forest fire occurrences in central China and identifies potential causal factors. This information is crucial for developing targeted fire prevention strategies and management recommendations to effectively safeguard forest resources and the ecological environment.

(a) Global Autocorrelation

(b) Local Autocorrelation

Figure 6. For autocorrelation analysis plots: (a) portrays global autocorrelation, while (b) demonstrates local autocorrelation.
3.3. Evaluation of Forecast Precision for Forest Fires in Central China

The metric calculations were performed using Python 3.5 software, and the accuracy evaluation chart was implemented in Excel 2020. As illustrated in Figure 7, in the training phase, the model achieved an impressive accuracy rate of 86.00%, precision of 88.00%, recall of 87.00%, and an F1 score of 87.50%. These metrics highlight the model’s exceptional and balanced performance in predicting forest fires. Furthermore, the AUC value of 90.50% underscores its outstanding ability to distinguish between fire occurrences and non-occurrences, demonstrating a high level of reliability in its predictive capabilities.

![Figure 7. Evaluation of the deep learning model’s effectiveness.](image)

The model’s effectiveness is further validated by its performance on the validation set. It achieved an accuracy of 85.00%, precision of 87.00%, recall of 86.00%, and an F1 score of 86.50%, along with an AUC value of 90.00%. These results collectively indicate that the model possesses a remarkable ability to generalize across various conditions, ensuring dependable and consistent predictions of forest fire incidents.

In summary, the model’s high accuracy, precision, recall, F1 score, and AUC values in both training and validation phases exemplify its proficiency in predicting forest fires. These robust performance metrics underscore its potential as a valuable tool for forest fire management, offering a solid foundation for informed decision making and strategic planning in fire prevention and control efforts.
3.4. Predicting Season Forest Fires in the Central China

Figure 8 provides a detailed overview of the seasonal patterns of fire-prone areas in central China, highlighting varying risk levels across different times of the year:

(i) Spring: During this season, specific regions in Hunan Province (Hengyang, Loudi, Yongzhou) and Hubei Province (Huangshi, Xianning) experience significantly elevated forest fire risks. The combination of rising temperatures, dry conditions with minimal rainfall, and accumulated dead vegetation from winter creates ideal fuel conditions. Agricultural practices like burning crop residues and increased tourist activities further contribute to the heightened risk. In Heilongjiang Province (Yichun, Heihe) and Jilin Province, frequent lightning strikes and geographical features such as steep mountains and dense vegetation also increase the likelihood of fires.

(ii) Summer: The onset of the rainy season generally reduces forest fire risks across most areas in the southern part of Central China due to increased humidity. However, regions like Huangshi in Hubei Province and parts of Heilongjiang Province (Yichun, Heihe) may still face higher risks due to uneven rainfall distribution or localized drought conditions.

(iii) Autumn: Forest fire risks increase again in autumn, particularly in places such as Binzhou, Yongzhou, and Hengyang in Hunan Province. As temperatures decrease and humidity levels drop, fallen leaves provide additional fuel for fires. Agricultural activities during the harvest season and increased tourism further elevate the risk during this period.

(iv) Winter: Despite colder temperatures, winter poses significant fire risks in regions such as Hengyang, Chenzhou, and Yongzhou in Hunan Province, as well as Huangshi in Hubei Province. Dry climates and low humidity contribute to the risk, exacerbated by dry vegetation and accumulated combustible materials. Human activities such as land clearing for agriculture also increase the likelihood of fires during winter.

Our seasonal forest fire risk predictions align consistently with monthly research findings from relevant scholars [18]. Seasonal strategies for forest fire prevention and control in central China vary significantly. During spring, heightened fire risks in regions like Hunan (e.g., Hengyang, Loudi, Yongzhou) and Hubei (e.g., Huangshi, Xianning) necessitate strict management of agricultural residues, outdoor fire bans, enhanced awareness among residents and tourists, and improved fire monitoring systems. In summer, despite generally lower fire risks, precautions against local droughts and uneven rainfall include maintaining firebreaks, controlled burns, monitoring dry lightning-prone areas, and swift fire response coordination between communities and authorities. Autumn sees increased fire risks in places such as Binzhou, Yongzhou, and Hengyang, requiring controlled burns of crop residues, strict supervision of outdoor activities, comprehensive fire risk assessments, and effective allocation of firefighting resources. Winter poses heightened risks due to dry conditions despite cold temperatures. Strategies include maintaining firefighting readiness, regular patrols in high-risk areas, utilizing satellite monitoring for early detection, and community engagement through fire prevention education and training. These strategies aim to protect forest resources and ensure community safety against the devastating impact of wildfires throughout the seasons in central China.
Figure 8. Seasonal zoning for forest fire predictions in southern China, where categories I through V denote the spectrum of forest fire occurrences from scarcely minimal to critically high levels.
4. Discussion and Conclusions

4.1. Discussion

In the central region of China, our study represents a significant advancement in the field of forest fire prediction modeling, employing state-of-the-art deep learning techniques, particularly Convolutional Neural Networks (CNNs), alongside a rich and detailed dataset. This research not only introduces a fresh perspective on understanding and forecasting forest fire but also deepens our insights into how various factors—such as climate dynamics, human activities, and geographical characteristics—affect the occurrence of forest fires.

The primary breakthrough of this study lies not only in confirming previous research findings [18,61,69] but also in markedly improving the accuracy of our predictive model. This enhancement provides a robust scientific foundation for shaping effective forest fire management strategies. By harnessing CNN algorithms, we surpassed the limitations of traditional statistical and physical models, significantly enhancing our ability to process and interpret vast amounts of complex and multidimensional data. CNNs excel in automatically extracting essential features from diverse datasets, thereby boosting the precision and reliability of our predictions.

Our innovative model holds immense promise for real-time forest fire monitoring and management. Through the integration of advanced technologies—such as satellite remote sensing, the BeiDou navigation system, and mobile communication networks—the model can continuously monitor critical fire-related parameters, like temperature, humidity, wind patterns, and vegetation health. This real-time monitoring capability empowers emergency response teams and decision makers to swiftly address fire incidents, bolstering proactive management strategies throughout China’s central region. Kernel density analysis provides a clear visual identification of these high-risk areas, offering crucial insights for relevant authorities to devise targeted fire prevention measures. For these identified high-risk zones, strategies such as intensified patrols, enhanced fire prevention infrastructure, and heightened public awareness campaigns can effectively reduce the likelihood of fire outbreaks. Furthermore, further research into fire occurrence patterns and underlying drivers in these regions will provide valuable scientific support for long-term forest fire prevention planning.

This study harnesses advanced deep learning, notably Convolutional Neural Networks (CNNs), to drastically elevate the accuracy of forest fire prediction in China’s central region, pinpointing high-risk zones. This robust foundation underpins refined, zone-specific management strategies. By precisely forecasting fire risks, we optimize resource allocation, streamline emergency response, and mitigate threats to life and property. Committed to safeguarding the region’s invaluable forest ecosystems, our research fosters ecological balance through scientific stewardship, ensures sustainable resource use, and fosters harmony between humans and nature. By identifying high-risk areas and implementing targeted prevention, our model safeguards China’s forests, balances ecology, and secures human well-being. Its potential transcends into early warning systems, resource optimization, and disaster coordination, positioning it as a cornerstone for sustainable coexistence with nature.

Looking ahead, our research agenda will continue to evolve. We are committed to exploring the impact of extreme climate phenomena—such as El Niño-Southern Oscillation (ENSO) and La Niña [87–90]—on forest fire dynamics. Additionally, we will refine our prediction model to further enhance accuracy, leveraging advancements in higher-resolution fire detection technologies [91,92]. Despite challenges related to acquiring consistent and precise data across expansive regions, we remain dedicated to overcoming these obstacles.

However, it is important to note a significant limitation in our study: the absence of comprehensive lightning data. Lightning strikes are a critical trigger for forest fires, yet our research did not directly incorporate data from lightning detection networks due to budgetary and logistical constraints. This omission underscores the complexity and cost associated with acquiring precise and comprehensive lightning data across large geographical areas, which are essential for accurately predicting forest fire. Moreover, our
forests will focus on distinguishing between different types of forest fires, including those initiated by lightning strikes [93,94] and human activities [95]. To more realistically simulate the dynamics of forest fires, we will explore more comprehensive and detailed datasets, integrating higher-resolution data and a broader array of fire descriptors. This will aid us in developing prediction models capable of distinguishing between different fire types [96–98], providing deeper insights and more effective strategies for fire prevention and management.

4.2. Conclusions

The exhaustive analysis conducted in this study offers profound insights into the dynamics of forest fires in central China, informing strategic management approaches. The following are the key findings and corresponding conclusions of our research:

Key Finding 1: Spatial Distribution Characteristics of Fires
Conclusion: Leveraging GIS technology, we identified significant spatial clustering and autocorrelation of fire occurrences, particularly in Hunan, Jiangxi, and the northeastern regions. This underscores the urgency for targeted prevention and management strategies to mitigate ecological, safety, and economic risks.

Key Finding 2: Predictive Performance of the CNN Model
Conclusion: The developed Convolutional Neural Network (CNN) model achieved remarkable predictive performance, with an accuracy of 86.00%, precision of 88.00%, recall of 87.00%, F1 score of 87.50%, and an AUC value of 90.50%. This demonstrates the value of integrating advanced machine learning with spatial analysis for proactive fire management, enabling precise decision making.

Key Finding 3: Seasonal Fire Risk Patterns
Conclusion: Seasonal analysis revealed distinct patterns, with spring and autumn posing heightened risks due to dry weather, agricultural activities, and human presence. Conversely, summer and winter exhibit localized threats, emphasizing the need for seasonally adapted fire prevention measures.

In summary, effective forest fire management in central China necessitates a harmonious blend of GIS-based spatial analysis and predictive modeling techniques, such as CNNs. By implementing targeted zoning strategies and leveraging the predictive prowess of the model, authorities can enhance preparedness, response, and mitigation efforts, thereby safeguarding ecosystems, protecting public safety, and maintaining economic stability in forested regions.

Author Contributions: Y.G., S.B. and Q.H. contributed to this study, encompassing research design, experimental execution, data analysis, and manuscript writing and revision. The expertise and collaborative spirit of each author were instrumental in ensuring the successful progression of the research and the achievement of results. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the International Cooperation and Exchange Program of the National Natural Science Foundation of China “Forest-Grassland Fire Risk Warning and Information Sharing on the Mongolian Plateau under the background of Climate Change” (41961144019) and the research project of the key research base of humanities and social sciences in universities of Inner Mongolia Autonomous Region—Yinshan Cultural Research Center of China (23YSYJ00014).

Data Availability Statement: Data are contained within the article.

Acknowledgments: We wish to extend our sincere appreciation to the editors and reviewers whose invaluable insights and recommendations significantly enriched the quality of this study.

Conflicts of Interest: The authors declare no conflicts of interest.
Appendix A

The main fire point, road and settlements data maps of this study are as follows.
Appendix B

<table>
<thead>
<tr>
<th>Convolutional Neural Network (CNN) Pseudocode Demonstration: Forest Fire Occurrence Prediction</th>
</tr>
</thead>
<tbody>
<tr>
<td>import numpy as np</td>
</tr>
<tr>
<td>import tensorflow as tf</td>
</tr>
<tr>
<td>from tensorflow.keras.models import Sequential</td>
</tr>
<tr>
<td>from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense, Dropout,</td>
</tr>
<tr>
<td>BatchNormalization, LeakyReLU</td>
</tr>
<tr>
<td>from tensorflow.keras.optimizers import SGD</td>
</tr>
<tr>
<td>from sklearn.metrics import accuracy_score, precision_score, recall_score, f1_score, roc_auc_score</td>
</tr>
<tr>
<td>from sklearn.model_selection import train_test_split</td>
</tr>
<tr>
<td># Assuming you have a function to load your dataset</td>
</tr>
<tr>
<td>def load_dataset():</td>
</tr>
<tr>
<td># Replace with your actual data loading and preprocessing steps</td>
</tr>
<tr>
<td>X, y = load_data()</td>
</tr>
<tr>
<td>X_train, X_test, y_train, y_test = train_test_split(X, y, test_size = 0.2, random_state = 42)</td>
</tr>
<tr>
<td>X_train, X_val, y_train, y_val = train_test_split(X_train, y_train, test_size = 0.2, random_state = 42)</td>
</tr>
<tr>
<td>return X_train, X_val, X_test, y_train, y_val, y_test</td>
</tr>
<tr>
<td># Step 1: Define the CNN model</td>
</tr>
<tr>
<td>def create_cnn_model(input_shape, num_classes):</td>
</tr>
<tr>
<td>model = Sequential()</td>
</tr>
<tr>
<td># Convolutional layers</td>
</tr>
<tr>
<td>model.add(Conv2D(64, (3, 3), activation = 'relu', input_shape = input_shape))</td>
</tr>
<tr>
<td>model.add(MaxPooling2D((2, 2)))</td>
</tr>
<tr>
<td>model.add(Conv2D(128, (3, 3), activation = 'relu'))</td>
</tr>
<tr>
<td>model.add(MaxPooling2D((2, 2)))</td>
</tr>
<tr>
<td>model.add(Conv2D(256, (3, 3), activation = 'relu'))</td>
</tr>
<tr>
<td>model.add(MaxPooling2D((2, 2)))</td>
</tr>
<tr>
<td>model.add(Flatten())</td>
</tr>
<tr>
<td># Fully connected layers</td>
</tr>
<tr>
<td>model.add(Dense(128, activation = 'relu'))</td>
</tr>
<tr>
<td>model.add(Dropout(0.5))</td>
</tr>
<tr>
<td># Output layer</td>
</tr>
<tr>
<td>model.add(Dense(num_classes, activation = 'sigmoid'))</td>
</tr>
<tr>
<td># Assuming binary classification</td>
</tr>
<tr>
<td># Compile the model</td>
</tr>
<tr>
<td>optimizer = SGD(learning_rate = 0.001, momentum = 0.9, decay = 0.01)</td>
</tr>
<tr>
<td>model.compile(optimizer = optimizer, loss = 'binary_crossentropy', metrics = ['accuracy'])</td>
</tr>
<tr>
<td>return model</td>
</tr>
<tr>
<td># Step 2: Load and preprocess the dataset</td>
</tr>
<tr>
<td>X_train, X_val, X_test, y_train, y_val = load_data()</td>
</tr>
<tr>
<td># Step 3: Create an instance of the CNN model</td>
</tr>
<tr>
<td>input_shape = X_train[0].shape</td>
</tr>
<tr>
<td>num_classes = 1</td>
</tr>
<tr>
<td>cnn_model = create_cnn_model(input_shape, num_classes)</td>
</tr>
<tr>
<td># Step 4: Train the model</td>
</tr>
<tr>
<td>cnn_model.fit(X_train, y_train, epochs = 10, batch_size = 32, validation_data = (X_val, y_val))</td>
</tr>
<tr>
<td># Step 5: Evaluate the model’s performance on the test set</td>
</tr>
<tr>
<td>y_pred = cnn_model.predict(X_test)</td>
</tr>
<tr>
<td>y_pred_binary = (y_pred &gt; 0.5).astype(int)</td>
</tr>
<tr>
<td>accuracy = accuracy_score(y_test, y_pred_binary)</td>
</tr>
<tr>
<td>precision = precision_score(y_test, y_pred_binary)</td>
</tr>
<tr>
<td>recall = recall_score(y_test, y_pred_binary)</td>
</tr>
<tr>
<td>f1 = f1_score(y_test, y_pred_binary)</td>
</tr>
<tr>
<td>auc = roc_auc_score(y_test, y_pred)</td>
</tr>
<tr>
<td># Print evaluation metrics</td>
</tr>
<tr>
<td>print(f&quot;Test Accuracy: {accuracy:.4f}&quot;&quot;)</td>
</tr>
<tr>
<td>print(f&quot;Precision: {precision:.4f}&quot;&quot;)</td>
</tr>
<tr>
<td>print(f&quot;Recall: {recall:.4f}&quot;&quot;)</td>
</tr>
<tr>
<td>print(f&quot;F1 Score: {f1:.4f}&quot;&quot;)</td>
</tr>
<tr>
<td>print(f&quot;AUC: {auc:.4f}&quot; )</td>
</tr>
</tbody>
</table>


40. Abid, F. A survey of machine learning algorithms based forest fires prediction and detection systems. *Fire Technol.* 2021, 57, 559–590. [CrossRef]


58. Coskuner, K.A. Assessing the performance of MODIS and VIIRS active fire products in the monitoring of wildfires: A case study in Turkey. *IForest-Biogeosciences For.* 2022, 15, 85. [CrossRef]


73. Monjáras-Vega, N.A.; Briones-Herrera, C.I.; Vega-Nieva, D.J.; Calleros-Flores, E.; Corral-Rivas, J.J.; López-Serrano, P.M.; Pompa-García, M.; Rodríguez-Trejo, D.A.; Carrillo-Parra, A.; González-Cabán, A. Predicting forest fire kernel density at multiple scales with geographically weighted regression in Mexico. *Sci. Total Environ.* 2020, 718, 137313. [CrossRef]


75. Liu, Z.; Ge, C.; Zheng, K.; Bao, S.; Cui, Y.; Yuan, Y.; Zhang, Y. Forecasting Daily PM2.5 Concentrations in Wuhu with a Spatial-Autocorrelation-based Long Short-Term Memory Model. *Atmos. Environ. 2024, 331, 120605. [CrossRef]


78. Yang, J.; Liu, Q.; Deng, M. Spatial hotspot detection in the presence of global spatial autocorrelation. *Int. J. Geogr. Inf. Sci. 2023, 37, 1787–1817. [CrossRef]


88. Bai, M.; Wang, X.; Yao, Q.; Fang, K. ENSO modulates interaction between forest insect and fire disturbances in China. Nat. Hazards Res. 2022, 2, 138–146. [CrossRef]
93. Jiao, Q.; Fan, M.; Tao, J.; Wang, W.; Liu, D.; Wang, P. Forest fire patterns and lightning-caused forest fire detection in Heilongjiang Province of China using satellite data. Fire 2023, 6, 166. [CrossRef]
98. Harvey, B.J. Human-caused climate change is now a key driver of forest fire activity in the western United States. Proc. Natl. Acad. Sci. USA 2016, 113, 11649–11650. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.