Article

Implementation of Large Language Models and Agricultural Knowledge Graphs for Efficient Plant Disease Detection

Xinyan Zhao [†]**, Baiyan Chen** [†]**, Mengxue Ji** [†]**, Xinyue Wang, Yuhan Yan, Jinming Zhang, Shiyingjie Liu, Muyang Ye and Chunli Lv** *

China Agricultural University, Beijing 100083, China
* Correspondence: lvcl@cau.edu.cn
[†] These authors contributed equally to this work.

Abstract: This study addresses the challenges of elaeagnus angustifolia disease detection in smart agriculture by developing a detection system that integrates advanced deep learning technologies, including Large Language Models (LLMs), Agricultural Knowledge Graphs (KGs), Graph Neural Networks (GNNs), representation learning, and neural-symbolic reasoning techniques. The system significantly enhances the accuracy and efficiency of disease detection through an innovative graph attention mechanism and optimized loss functions. Experimental results demonstrate that this system significantly outperforms traditional methods across key metrics such as precision, recall, and accuracy, with the graph attention mechanism excelling in all aspects, particularly achieving a precision of 0.94, a recall of 0.92, and an accuracy of 0.93. Furthermore, comparative experiments with various loss functions further validate the effectiveness of the graph attention loss mechanism in enhancing model performance. This research not only advances the application of deep learning in agricultural disease detection theoretically but also provides robust technological tools for disease management and decision support in actual agricultural production, showcasing broad application prospects and profound practical value.

Keywords: plant disease detection; agricultural knowledge graphs; agricultural large model; deep learning; smart agriculture

1. Introduction

With the rapid development of smart agriculture, disease detection technology plays a crucial role in improving crop yield and quality [1,2]. However, traditional disease detection methods often rely on experienced experts or complex chemical tests, which are not only costly but also inefficient [3]. The early and accurate identification of crop diseases is essential for timely intervention, preventing significant yield losses and ensuring the quality of agricultural products [4].

Traditional methods of disease detection typically involve visual inspection by experts or laboratory-based chemical analysis. Although these methods may be effective, they are usually time-consuming, require substantial expertise, and are not scalable across large agricultural areas. Moreover, reliance on manual inspection means that disease detection can be severely delayed, allowing the disease to spread before appropriate measures are taken. Techniques such as calculating canopy temperature and vegetation indices from thermal imaging and hyperspectral imaging, analyzing early disease through linear and radial basis kernel of Linear Discriminant Analysis and Support Vector Machine classification methods, and applying these vegetation index combinations for quantifying and differentiating the severity of red leaf spots have been explored [5]. Near-Infrared (NIR) spectroscopy, based on the absorption of electromagnetic radiation at NIR wavelengths, is widely used for classifying and detecting chemical properties in grains and nuts, facilitating quality and process control [6].

The advent of smart agriculture has introduced the potential to automate and enhance the disease detection process using advanced technologies such as artificial intelligence and machine learning. Savarimuthu, Nickolas et al. [7] explored the potential of computer vision-based object detection methods for the early detection of plant diseases, conducting comparative studies using three different benchmark object detection models: YOLOv4, EfficientDet, and Scaled-YOLOV4. The findings indicate that the performance of EfficientDet is inferior to that of Scaled-YOLOV4. Wang et al. [8] developed an optimized lightweight YOLOv5 model for plant disease detection and classification. The optimized model demonstrated an 11.8% improvement in runtime and a 3.98% increase in accuracy compared to the original model, with an F1 score reaching 92.65%. Building on this, Qadri Syed Asif Ahmad et al. [9] employed YOLOv8 for plant leaf disease detection and segmentation. As the model is trained end to end, it effectively learns and generalizes from input data, thereby enhancing its predictive performance on unseen or novel instances of leaf diseases. The results showcase strong performance in accurately detecting and segmenting diseased areas. Malik Muneeb Elahi et al. [10] proposed a robust and improved method for weed detection using YOLOv9. The results indicate that YOLOv9 outperforms other deep learning models, with a performance increase of 2.15% over YOLOv8. Sankareshwaran, Senthil Pandi et al. [11] introduced a novel rice disease detection method termed the Cross-enhanced Artificial Hummingbird Algorithm-based AX-RetinaNet (CAHA-AXRNet). Compared to other existing rice disease detection methods, the proposed CAHA-AXRNet method achieves an accuracy of 98.1%. Zhang et al. [12] introduced a lightweight visual segmentation model named TinySegformer, specifically designed for agricultural pest detection, achieving real-time image processing at 32.7 frames per second on edge devices.

However, many existing AI-based solutions are limited by the quality and comprehensiveness of training data. Large Language Models (LLMs) such as ChatGPT and Bard have revolutionized natural language understanding and generation [13]. With their deep language comprehension, human-like text generation capabilities, context awareness, and strong problem-solving abilities, they hold invaluable value across various domains, such as search engines, customer support, and translation [14]. These models have achieved significant results in the fields of natural language processing and computer vision. Their robust feature extraction and information processing capabilities can be applied to various tasks in the agricultural domain. For instance, Osinga et al. explored from four perspectives: focusing on identifying the transformative drivers behind recognition applications; studying the big data characteristics of the research problem—quantity, speed, variety, and veracity; assessing the maturity of anticipated solutions; and paying attention to the problem domain or solution provider. This indicated that the segmentation within the agricultural domain itself hindered the broad exploration of various advancements in big data and artificial intelligence [15].

Agricultural Knowledge Graphs (KGs) are structured representations of agricultural information [16], covering a wide range of data, including crop types, disease symptoms, environmental factors, and management practices. These graphs can model complex relationships between different entities within the agricultural domain. By combining these knowledge graphs [17] with LLMs, systems can be created that not only understand the agricultural terminology and context but also use this understanding to improve disease detection and decision-making [18]. Peng et al. explored using domain-agnostic general pretrained LLMs to extract structured data from agricultural documents with minimal or no human intervention, building an LLM-based retrieval that efficiently extracts agricultural information from unstructured data [19]. Additionally, the working principle of the graph attention mechanism is to assign different weights to different nodes in the knowledge graph based on their relevance to the current task [20]. This dynamic adjustment ensures that the model pays more attention to the most relevant information, particularly useful in the context of disease detection, as certain symptoms or environmental factors may better illustrate the issue. Zhou et al. constructed an "Image–Text" multi-modal cooperative representation and knowledge-assisted disease identification model (ITK-Net), achieving

the highest accuracy, precision, sensitivity, and specificity rates at 99.63%, 99%, 99.07%, and 99.78%, respectively [21].

To address these issues, this paper proposes an innovative solution: enhancing disease identification and decision-support systems by integrating Agricultural Knowledge Graphs and LLM [22]. Initially, an LLM based on Agricultural Knowledge Graphs was constructed, which not only incorporated a vast amount of agricultural expertise but also optimized the extraction and utilization of information through natural language processing technology. Subsequently, a graph attention-based agricultural LLM structure was introduced, dynamically adjusting the flow of information within the network to better focus on key features related to disease identification, thus enhancing detection accuracy and robustness [23].

Moreover, a graph loss function was developed, an innovative loss function design aimed at optimizing the model training process by precisely adjusting the penalties for incorrect predictions, further enhancing the model's ability to recognize complex disease patterns. Through the integrated application of these technologies, a elaeagnus angustifolia disease detection model and the corresponding smart agriculture system were successfully developed. This system not only effectively identifies and classifies various elaeagnus angustifolia diseases but also provides operational recommendations, aiding agricultural producers in making more accurate agricultural management decisions.

In summary, this research not only innovates technologically but also provides robust support and fresh perspectives for the practical application of elaeagnus angustifolia disease detection and smart agriculture. By implementing the fusion of these advanced technologies, it is anticipated that smart agriculture will be propelled towards more efficient, precise, and intelligent directions.

2. Related Work

2.1. Knowledge Graphs

KGs, as effective tools for knowledge management [17], organize and integrate a vast array of agricultural data structurally [24], including types of crops, growth cycles, and disease information. By constructing KGs, entities and their relationships are defined within the graph, thus forming a multi-dimensional and highly connected information network. In the application of agricultural disease detection, KGs [25] not only aid researchers in quickly understanding complex relationships between crop diseases but also support advanced reasoning and querying capabilities, enhancing the accuracy of disease prediction and prevention strategies. For instance, by analyzing pathways and patterns within the graph [26], risks of diseases under specific environmental conditions can be predicted, allowing for preemptive measures to be taken.

KGs hold significant application value in the detection and management of agricultural diseases [27]. By structurally representing information about different diseases, including characteristics, influencing factors, transmission methods, and management approaches within the knowledge graph, the rapid diagnosis and precise control of diseases can be achieved [28]. For example, potential risks of diseases under specific environmental conditions can be forecasted by analyzing pathways and patterns within the graph, facilitating preventive actions. In crop management, KGs can integrate various factors affecting crop growth, such as climatic conditions, soil types, fertilization plans, and irrigation strategies, thereby providing comprehensive guidance for crop management [29]. For instance, by analyzing the knowledge graph, optimal combinations of nutrients required at different growth stages of a crop can be identified, optimizing fertilization plans and enhancing crop yield and quality [30].

Furthermore, KGs are also utilized in monitoring and predicting crop growth conditions [31]. By comparing real-time meteorological and sensor data with historical data within the knowledge graph, trends in crop growth and potential issues can be predicted, assisting farmers in timely adjusting management strategies to avoid or minimize losses.

2.2. Large Language Models

On the other hand, LLMs such as GPT and BERT have demonstrated outstanding performance in the field of natural language processing; their training involves massive textual data, enabling the capture of deep linguistic semantics and complex structures [32]. The application of LLMs to agricultural disease detection allows for the powerful text comprehension capabilities of the model to be utilized in analyzing and processing unstructured data such as agricultural literature, research reports, and field records [33]. For instance, information regarding disease characteristics, influencing factors, and management methods can be automatically extracted from the literature, providing data support and decision-making recommendations for disease management. Moreover, by integrating with KGs, LLMs can more precisely understand and generate queries and outputs related to agricultural knowledge, enhancing the overall intelligence level of the system.

The architecture of LLMs is as follows [34]: the Transformer architecture forms the foundation of LLMs, primarily comprising encoders and decoders. Inputs are first transformed into word embeddings. Assuming the input sequence is $X = (x_1, x_2, ..., x_n)$, each x is converted into a corresponding embedding vector e. Attention scores are computed as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \tag{1}$$

where the query (Q), key (K), and value (V) are represented, and d_k is the dimension of the key vector, with $\frac{1}{\sqrt{d_k}}$ acting as the scaling factor. To capture different attention patterns, the Transformer utilizes a multi-head attention mechanism, dividing the input into multiple heads, computing attention separately, and then concatenating the results:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, ..., \text{head}_h)W_O \tag{2}$$

Each head computes attention similarly to the single-head attention mechanism:

$$\text{head}_i = \text{Attention}(QW_Q^i, KW_K^i, VW_V^i) \tag{3}$$

As Transformers lack convolutional and recursive structures and cannot directly handle sequential information, position encoding is added:

$$\text{PE}_{(pos,2i)} = \sin\left(\frac{pos}{10000^{2i/d_{model}}}\right), \quad \text{PE}_{(pos,2i+1)} = \cos\left(\frac{pos}{10000^{2i/d_{model}}}\right) \tag{4}$$

Each attention layer is followed by a feed-forward network:

$$\text{FFN}(x) = \text{ReLU}(xW_1 + b_1)W_2 + b_2 \tag{5}$$

Layer normalization and residual connections follow each sublayer (self-attention mechanism and feed-forward network):

$$\text{LayerNorm}(x + \text{Sublayer}(x)) \tag{6}$$

LLMs like BERT and GPT are initially pretrained on large-scale unlabeled data to capture general features. BERT's pretraining task is the Masked Language Model (MLM), which randomly masks some words in the input sequence and prompts the model to predict these masked words. GPT's pretraining task is an Autoregressive Language Model—predicting the next word based on the previous text. After pretraining, the model is fine-tuned on specific task datasets. The fine-tuning process involves further training the model on labeled data to adapt it to specific tasks, such as text classification or question-answering systems.

LLMs can serve as part of agricultural decision-support systems, offering scientifically based decision recommendations. For example, in the selection of crop planting, LLMs can recommend the most suitable crop varieties based on factors such as climatic conditions, soil types, and market demands [35]. By analyzing historical data and current environmental

conditions, LLMs can also provide planting plans and yield predictions, helping farmers formulate rational production strategies [36]. In the field of disease prevention and control, LLMs can analyze the transmission paths and influencing factors of diseases, providing effective prevention and control schemes [37]. For instance, based on disease characteristics and environmental conditions, LLMs can recommend appropriate pesticides and control measures, assisting farmers in reducing disease losses and enhancing crop yield and quality [38].

LLMs can also be applied to agricultural intelligent question-answering systems, providing convenient knowledge services to farmers and agricultural practitioners [39]. Through natural language processing technology, users can pose questions to the system in natural language, which are understood and generated by LLMs to provide accurate answers [40]. For example, a farmer may inquire about the symptoms and prevention methods of a particular disease, and the system, utilizing the knowledge within LLMs, will provide detailed responses and recommendations. Agricultural intelligent question-answering systems not only address common queries but also offer personalized advice. For instance, users can input information about their farms and production situations, and the system, based on this information, provides customized planting and management advice, helping farmers improve production efficiency and profits [41].

2.3. Analysis of KG and LLM

KGs are employed as effective knowledge management tools across various industries. For instance, in the healthcare sector, KGs are utilized to integrate patient information, disease diagnostics, and treatment plans, supporting advanced reasoning and querying capabilities within medical decision systems. In the financial services industry, KGs are applied to risk assessment, customer relationship management, and fraud detection. In contrast, the application of KGs within this study, especially in smart agriculture, particularly in disease identification, crop management, and production decision support, showcases unique value. By constructing entities and relationships within the agricultural domain, KGs not only accelerate the processing of agricultural data but also enhance the precision of disease prevention and management.

Furthermore, LLMs such as GPT and BERT exhibit exceptional capabilities in text understanding and generation within the field of natural language processing. Compared to their applications in other industries—such as automated customer service systems and legal document analysis—LLMs in this study focus more on analyzing and processing unstructured data such as agricultural literature, research reports, and field records. Coupled with KGs, LLMs in this study are able to precisely understand and generate queries and outputs related to agricultural knowledge, which significantly elevates the overall level of system intelligence.

The primary distinction between KGs and LLMs in this study and related research in other industries or smart agriculture lies in their deep integration and complementary application. The combination of structured knowledge management from KGs and the powerful text-processing abilities of LLMs provides a novel solution for smart agriculture, capable of more comprehensively and accurately addressing complex agricultural issues. This integrated application not only optimizes data processing workflows but also enhances the adaptability and effectiveness of the models in practical applications.

3. Materials and Method

3.1. Dataset Collection

The collection of data plays a pivotal role in the research and development of the elaeagnus angustifolia disease detection model and smart agriculture system. This step is not only foundational for model training but also crucial for ensuring the accuracy and effectiveness of the model. To establish a comprehensive and precise training database, we focused on collecting image data covering five major elaeagnus angustifolia pests, with the number of images for each type of pest ranging from 800 to 1500, from 2023 to

2024. These pests include black leaf beetles, longhorn beetles, psyllids, scale insects, and mealybugs, which are the most common and damaging pests to elaeagnus angustifolia trees. Image collection was primarily conducted under natural lighting conditions to simulate the actual agricultural production environment as closely as possible. Furthermore, to capture the different stages of pest damage to plants, photographs were taken at various times, meticulously documenting the activities of the pests and their specific impacts on the plants.

During the photography process, high-resolution camera equipment was utilized, and camera settings such as aperture, shutter speed, and ISO were consistently maintained to ensure uniformity in the image quality. Prior to each shooting session, the environment was standardized, including ensuring a clean background and suitable lighting to minimize the impact of environmental factors on image quality. Through this rigorous and standardized method of image collection, a high-quality and representative dataset was obtained, providing a solid foundation for the training and evaluation of the model. The specific quantities and samples for each pest are shown in Table 1 and Figure 1.

Table 1. Composition of dataset.

Pest	Quantity
Longhorn beetle	890
Scale insect	983
Psyllid	1306
Black leaf beetle	1005
Mealybug	1492

Figure 1. Visual representation of the dataset. (**a**) Longhorn beetle, (**b**) scale insect, (**c**) psyllid, (**d**) black leaf beetle, and (**e**) mealybug.

Black leaf beetles are highly destructive pests commonly found on elaeagnus angustifolia trees, particularly active from early spring to summer. They damage the leaves extensively by chewing on the edges and surfaces. To study and detect such pests thoroughly, we collected 1005 representative images of elaeagnus angustifolia leaves damaged by black leaf beetles. These images meticulously record various stages of damage from mild to severe, providing rich data to support the model's learning to recognize different levels of pest damage. Longhorn beetles, primarily attacking the trunks and branches of elaeagnus angustifolia trees, burrow inside the wood during their larval stage, causing severe internal structural damage and even death of the trees over time. We prepared 890 images for this pest, showing the traces of adult and larval activities on different parts of the trees, along with the physical damage they cause. Psyllids, surviving by sucking plant sap, lead to the yellowing and wilting of elaeagnus angustifolia leaves, severely affecting the overall health and growth of the trees. For this, we collected 1306 images of psyllid damage, which not only record the activities of psyllids on the leaves but also include the symptoms caused by them, providing data to help the model learn to recognize specific diseases caused by psyllids. Scale insects often produce sticky secretions on the underside of leaves, causing sooty mold that further hinders photosynthesis, affecting the plant's growth and development. During the data collection, we paid special attention to the activity patterns of these pests and the specific symptoms they cause, collecting 983 images of scale insect damage. Finally, mealybugs, which form a protective hard shell, make

conventional physical and chemical control methods ineffective. We collected 1492 images showing the distribution, breeding patterns, and typical symptoms and damage caused by mealybugs on elaeagnus angustifolia trees.

Through this thorough data collection, we not only provide rich training materials for the elaeagnus angustifolia disease detection model but also ensure the model's high generalizability and robustness under real-world conditions. These image data cover different developmental stages of pests and include as many scenarios as possible under various lighting conditions, different shooting angles, and different background environments to simulate the complex conditions of the real world. In the future, this valuable data will be used to validate and test our proposed model structure based on Agricultural Knowledge Graphs–LLM and graph attention, and to compare it with existing deep learning models to assess their performance and effects in practical applications. We anticipate that through this series of research and development activities, we will ultimately achieve an efficient and accurate elaeagnus angustifolia disease detection smart system, providing solid technical support for the advancement of smart agriculture.

3.2. Dataset Preprocessing

3.2.1. Image Enhancement Based on Traditional Computer Vision Methods

In this study, image preprocessing is identified as a key step in enhancing the performance of the elaeagnus angustifolia disease detection model. Traditional computer vision methods play a fundamental and significant role in image enhancement by improving the visual quality of images, thereby enhancing the subsequent algorithms' ability to recognize diseases. In agricultural image processing, issues such as low contrast and high noise levels often arise due to external environmental influences like lighting and weather conditions, severely impacting the accuracy of disease detection.

The impact of lighting conditions on the model is primarily manifested in terms of image brightness, contrast, and color saturation. Improper lighting can cause images to be overexposed or underexposed, thereby affecting the model's ability to recognize pests and diseases. To adapt to varying lighting environments, image enhancement techniques were employed in this study to adjust the image brightness and contrast. These methods include histogram equalization and gamma correction. Histogram equalization enhances image contrast by expanding the brightness distribution, with the formula given by:

$$I'(x,y) = \frac{L-1}{MN} \sum_{i=0}^{I(x,y)} h(i) \tag{7}$$

where $I(x,y)$ is the brightness value at coordinates (x,y), $I'(x,y)$ is the enhanced brightness value, $h(i)$ is the frequency of brightness value i, M and N are the width and height of the image, and L is the number of possible pixel levels. Variations in shooting angles may affect how pests and diseases are represented in images, making them difficult for the model to recognize. To address this, images of pests and diseases were captured from multiple angles during data collection, and geometric transformations such as rotation and scaling were applied in the preprocessing phase to enhance the model's adaptability to images from different angles. The specific transformation is as follows:

$$I'(x',y') = I(x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \tag{8}$$

where θ is the rotation angle and (x',y') are the coordinates after rotation. The background complexity is another significant factor affecting model performance. To mitigate the impact of background noise, image segmentation techniques were used in the preprocessing stage to automatically separate the foreground from the background, analyzing only areas containing the target pests or diseases. This approach not only speeds up image processing

but also enhances the accuracy of the model. Common methods of image segmentation include threshold segmentation and region-growing algorithms, with the formula given:

$$S(x,y) = \begin{cases} 1 & \text{if } I(x,y) > T \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

where $S(x,y)$ is the segmented image, and T is the chosen threshold.

3.2.2. MixUp, CutMix, Copy

Advanced data augmentation techniques including MixUp, CutMix, and Copy are utilized to enhance the generalization ability and robustness of the elaeagnus angustifolia disease detection model as shown in Figure 2. These techniques introduce greater variability into the training data, helping the model learn a broader range of features, thus performing with higher accuracy and stability when faced with the variable real-world agricultural environment.
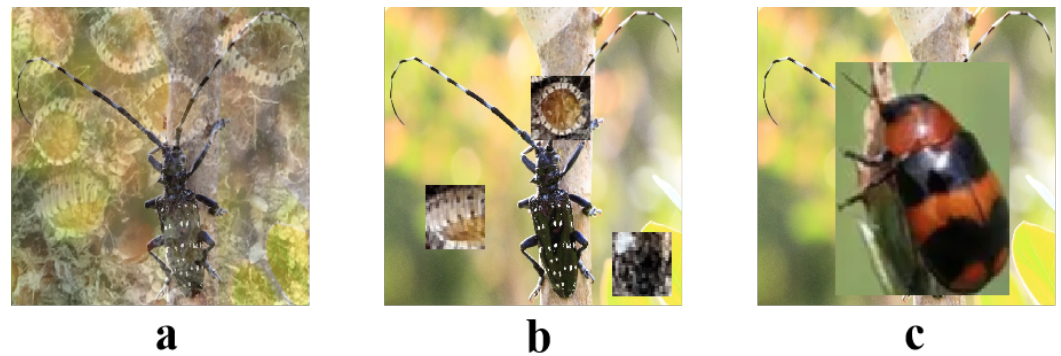

Figure 2. Image dataset enhancement method used in this paper. (**a**) MixUp, (**b**) Cutmix, and (**c**) Copy.

The MixUp method, a technique for data augmentation at the image level, generates new training samples by linearly interpolating the pixel values and their labels from two different images. Specifically, given two training samples (x_i, y_i) and (x_j, y_j), where x represents image data and y represents the corresponding labels, the new sample (x', y') generated by MixUp can be calculated using the following formulas:

$$x' = \lambda x_i + (1 - \lambda)x_j \tag{10}$$

$$y' = \lambda y_i + (1 - \lambda)y_j \tag{11}$$

where λ is a value randomly drawn from a Beta distribution $Beta(\alpha, \alpha)$, with α being a hyperparameter typically set between 0.2 and 0.4. This approach enables the model to learn features transitioning smoothly from one disease type to another, enhancing its capability to handle images with partially unclear labels. CutMix, another data augmentation technique, differs from MixUp's full-pixel blending by cutting out a rectangular area from one image and pasting it onto another, while correspondingly mixing the labels of these two images. The operation is performed as follows:

$$x' = M \odot x_i + (1 - M) \odot x_j \tag{12}$$

$$y' = \lambda y_i + (1 - \lambda)y_j \tag{13}$$

where M is a binary mask representing the area cut from image x_i, and λ is typically determined by the proportion of the area cut, i.e., $\lambda = \frac{1}{S} \sum M$, with S being the total number of pixels in the image. This method is particularly suitable for handling images with diverse object sizes and locations, such as different types and stages of elaeagnus

angustifolia diseases. The Copy technique, a simple copy-and-paste method, allows a part of the image (usually an object of interest, such as a diseased area) to be directly copied and pasted into another image. This method can artificially increase the frequency of target objects in images, thus aiding the model in better learning and recognizing these objects:

$$x' = x_i \text{ with patches from } x_j \tag{14}$$

Although straightforward, this method is highly effective in situations where training data are insufficient or certain classes of samples are rare. By applying these data augmentation techniques during the training process of the elaeagnus angustifolia disease detection model, significant improvements in model data coverage and robustness are achieved. These techniques not only adapt the model to diverse inputs but also enhance its ability to recognize anomalies and atypical disease manifestations. The combined use of MixUp, CutMix, and Copy provides strong data support for the development of the elaeagnus angustifolia disease detection model and smart agriculture system, effectively enhancing model performance.

3.3. Proposed Method

Overall

In this research, a comprehensive elaeagnus angustifolia disease detection model and smart agriculture system are proposed, integrating multiple advanced technological modules, including LLMs, KGs, Graph Neural Networks (GNNs), representation learning, neural-symbolic reasoning, and few-shot learning, as shown in Figure 3. These modules collaboratively form an intelligent system capable of processing complex agricultural data and providing accurate disease diagnostics. The construction process of the model and the interconnections between these modules are detailed below.
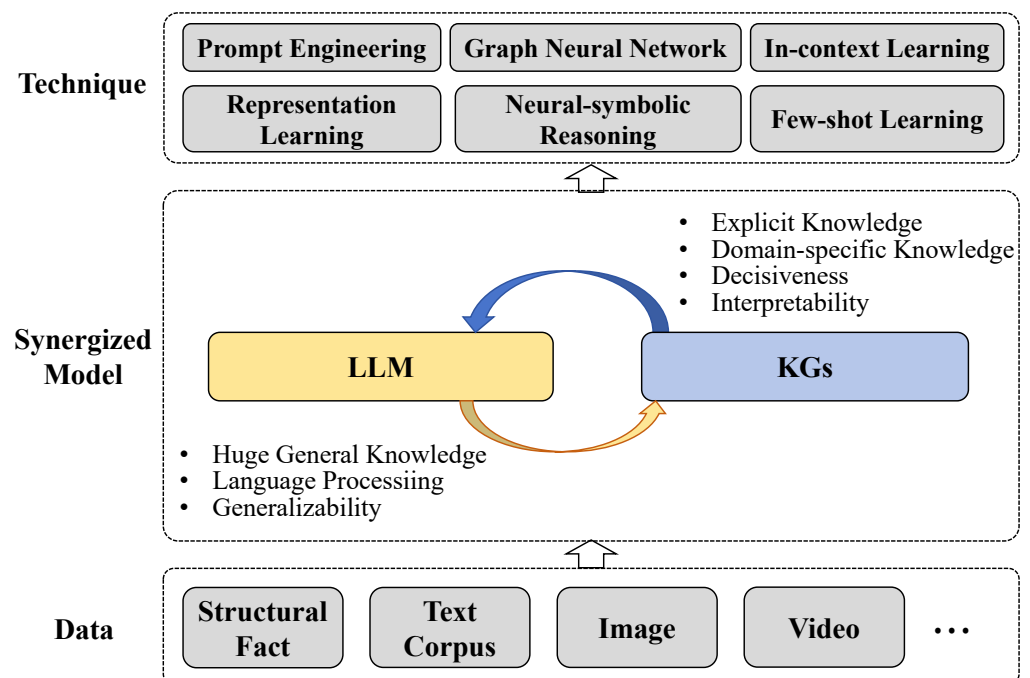

Figure 3. This figure presents the overall architecture of the elaeagnus angustifolia disease detection and smart agriculture system proposed in this paper. The system integrates a variety of advanced technologies, including LLM, Agricultural KGs, GNN, representation learning, neural-symbolic reasoning, and few-shot learning.

Initially, preprocessed image and text data are input into the system. For image data, feature extraction is performed using GNNs, which are particularly suited to handling graph-structured data. By leveraging connections between nodes, GNNs effectively learn

representations of node features. In the context of elaeagnus angustifolia disease detection, each image is considered a node, with connections potentially based on similarities or geographical locations between images. The state of the nodes is updated using the following formula:

$$h_v^{(l+1)} = \text{ReLU}\left(b^{(l)} + \sum_{u \in N(v)} \frac{1}{c_{uv}} W^{(l)} h_u^{(l)}\right) \tag{15}$$

where $h_v^{(l)}$ represents the feature representation of node v at layer l, $N(v)$ denotes the set of neighboring nodes of v, c_{uv} is the normalization constant between nodes u and v, $W^{(l)}$ and $b^{(l)}$ are the weight and bias at layer l respectively, and ReLU is the activation function. Simultaneously, text data are fed into a LLM. LLMs excel in natural language processing tasks, capable of understanding and generating specialized knowledge about elaeagnus angustifolia diseases. Additionally, the input and output of the model are optimized through Prompt Engineering technology to suit specific task requirements. Prompt Engineering is a technique that uses predefined prompts to guide model outputs, significantly enhancing model performance on specific tasks. After feature extraction, these features are directed to the representation learning module. This module is responsible for integrating features from both GNN and LLM, forming a unified representation space that allows the model to make decisions based on both image and text information. During the representation learning process, techniques of neural-symbolic reasoning and few-shot learning are employed to further optimize the model's learning efficiency and reasoning capabilities. Neural-symbolic reasoning combines the representational power of deep learning with the reasoning capabilities of symbolic logic, performing reasoning as follows:

$$s = \sigma\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{16}$$

where x_i are the input features, w_i are the weights, b is the bias, σ is the activation function, and s is the symbolic logic result derived from reasoning. Few-shot learning enables the model to quickly adapt to new tasks with very few samples, enhancing performance under conditions of scarce data. Through the above processes, the modules work in coordination not only to improve the accuracy of elaeagnus angustifolia disease detection but also to enhance the model's understanding and utilization of agricultural domain knowledge. This integrated approach provides a new technological pathway for the development of smart agriculture, showcasing the potential application of deep learning and natural language processing technologies in agricultural disease management.

3.4. Few-Shot Learning

Few-shot learning is a machine learning technique designed to address data scarcity, enabling models to quickly adapt to new tasks with limited training samples, thereby significantly enhancing model generalization capabilities. In traditional deep learning models, especially in fields like image recognition and natural language processing, a substantial amount of annotated data are usually required to train a stable and high-performing model. However, in practical applications, particularly in specific scenarios like agricultural disease detection, acquiring a large quantity of high-quality annotated data is often impractical due to the high specialization and labeling costs. Therefore, the introduction of few-shot learning is particularly crucial for our research, as it allows models to effectively learn and predict with only a few annotated samples. In our study, the application of few-shot learning focuses on several key areas:

1. Model Pretraining: We employ a pretraining and fine-tuning strategy, initially pretraining our model on a large-scale generic dataset to learn rich feature representations. During pretraining, the model captures basic, universal visual and graph structural features, laying the groundwork for subsequent few-shot learning.

2. Meta-Learning Strategy: For the implementation of few-shot learning, we use meta-learning techniques, specifically the Model-Agnostic Meta-Learning (MAML) algorithm. This approach involves training the model on multiple tasks, each with only a few samples. MAML seeks a good parameter initialization that allows the model to rapidly adapt to new tasks with minimal gradient updates, greatly enhancing the model's ability to quickly adapt to new disease types.

3. Data Augmentation: To maximize the use of limited samples, we also incorporate robust data augmentation techniques during data preprocessing, such as rotation, flipping, and scaling, to artificially expand the dataset. These techniques generate more variations from the existing few samples, increasing data diversity during model training and thus enhancing the model's generalization ability to unseen samples.

4. Embedding Learning: In the model structure design, we introduce an embedding learning mechanism, which involves learning to map input data into an embedding space where samples of the same category are closer together, and samples of different categories are farther apart. This strategy is particularly effective with only a few annotated samples, as it emphasizes the relative relationships between samples rather than absolute label information.

Through these strategies, our model achieves high recognition accuracy and good generalization performance in the specific field of jujube disease detection, even with relatively scarce training samples. Future research will continue to explore more efficient few-shot learning methods to further enhance the practicality and accuracy of the model in real agricultural applications.

3.4.1. Agricultural Knowledge Graph–Large Language Model (KG-LLM)

In this research, a comprehensive system combining an Agricultural KG with a LLM has been designed and implemented to enhance the accuracy and efficiency of elaeagnus angustifolia disease detection. The core of this system's design lies in the deep integration of domain-specific knowledge with advanced natural language processing technology, ensuring that the system can provide scientifically valid diagnostic support in practical applications as shown in Figure 4.
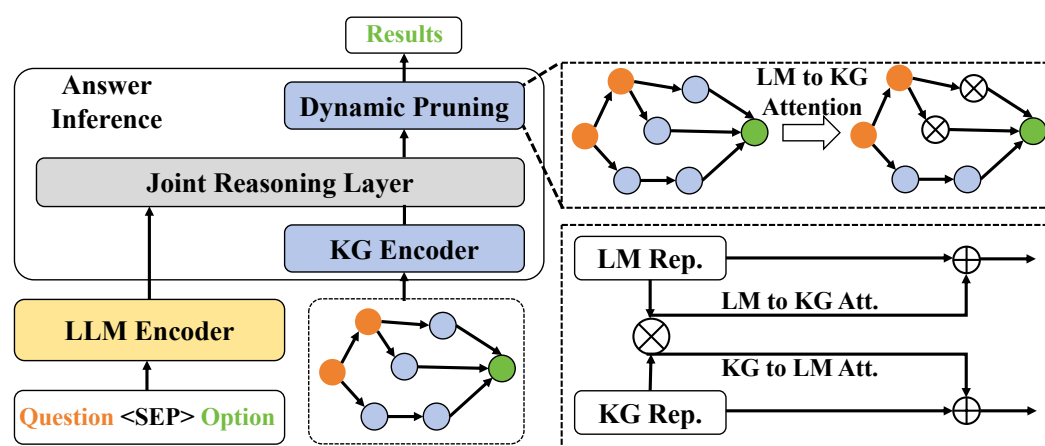

Figure 4. This figure displays the system architecture that combines Agricultural KGs and Large Language Models. It details the workflow and interaction between the Agricultural Knowledge Graph Encoder (KG Encoder) and the Large Language Model Encoder (LLM Encoder).

The system design comprises two main modules: the Agricultural KG Encoder and the LLM Encoder. The KG Encoder specifically processes structured knowledge related to elaeagnus angustifolia diseases, such as types of diseases, symptom descriptions, influencing factors, and preventive measures. This knowledge is organized in the form of a graph, where each node represents a knowledge entity, and edges represent relationships between entities. The LLM Encoder processes natural language data, extracting textual

information related to elaeagnus angustifolia diseases and generating text content that can be used for decision support. In implementation, a deep learning network is used to jointly train these two modules. The network structure is designed as follows: initially, the KG Encoder utilizes a GNN to extract features of entities and relationships, with a Multi-Layer Perceptron (MLP) encoding the information of each node. For the LLM Encoder, a pretrained Transformer model equipped with multiple layers of self-attention mechanisms is used, which captures long-range dependencies and enhances semantic understanding capabilities. The interactivity of the system is designed around a Dynamic Attention mechanism between the two modules. Attention from LLM to KG (LM to KG Attention) and from KG to LLM (KG to LM Attention) are introduced, allowing the system to dynamically adjust the flow and focus of information based on the current task. For instance, when generating a disease diagnostic report, the model can enhance focus on symptoms and preventive measures.

Moreover, to further enhance the system's decision-making capability, a Joint Reasoning Layer is added after the two encoders. This layer's task is to integrate information from both KG and LLM, producing the final decision output. In terms of design parameters, the network includes multiple convolutional layers with varying widths and heights, with channel numbers ranging from 64 to 256, ensuring sufficient model capacity to handle complex agricultural data. Such design not only makes our system innovative in theory but also effectively supports decision-making in practice, particularly in addressing the complex task of elaeagnus angustifolia disease detection, significantly improving diagnostic accuracy and efficiency. By deeply integrating structured KGs and unstructured textual data, the system provides users with a comprehensive, precise, and interactive decision-support tool, showcasing the broad application potential of artificial intelligence in smart agriculture.

3.4.2. Text–Knowledge Alignment Module

In the elaeagnus angustifolia disease detection and smart agriculture system, the text–knowledge alignment module plays a crucial role. The main task of this module is to ensure effective alignment between features recognized from images and textual information in the knowledge graph, thereby enhancing the system's diagnostic accuracy and reliability for elaeagnus angustifolia diseases as shown in Figure 5.
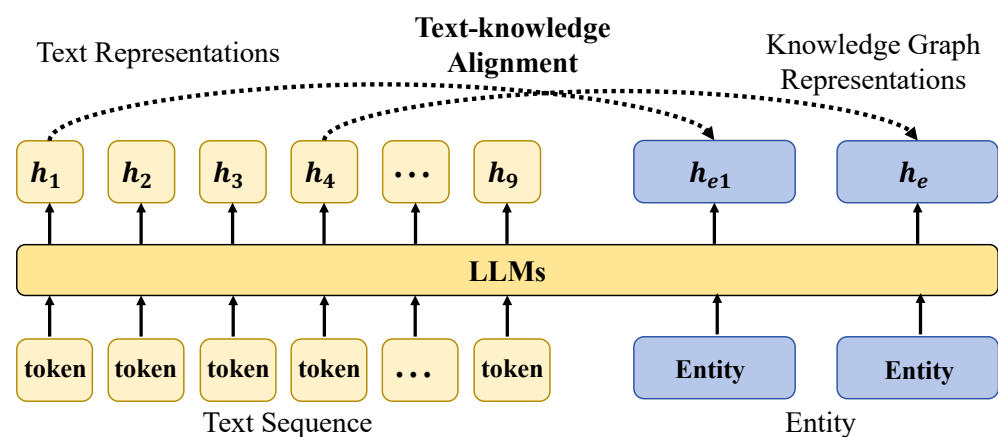

Figure 5. This figure demonstrates the working principle of the text–knowledge alignment module. The key aspect of this module is the effective alignment of text representations from natural language processing with entity representations from the knowledge graph. The figure elaborately shows the process of extracting feature vectors from text sequences and how these vectors are aligned with the entity vectors from the knowledge graph.

The core of this module involves processing and aligning text sequences with representations from the knowledge graph using deep learning technology. The implementation

consists of two parts: the text representation part and the knowledge graph representation part.

Text Representation Part: This section utilizes LLMs, such as BERT or GPT, to encode text data. By encoding text related to descriptions, reports, or literature on elaeagnus angustifolia diseases, the model generates a series of vector representations, each corresponding to a token in the text. These vectors capture the semantics and contextual information of the vocabulary, providing a basis for subsequent alignment operations. For example, the text sequence "elaeagnus angustifolia leaves show yellow spots, possibly due to iron deficiency symptoms." is transformed into a series of vectors $h_1, h_2, \ldots, h_n$.

In the alignment operation, an attention mechanism is employed to align the text vectors with the knowledge graph vectors. This mechanism dynamically selects the knowledge graph entities with the strongest corresponding relationships by calculating the similarity between text vectors and knowledge graph vectors, thus optimizing the system's information integration.

In designing the text–knowledge alignment module, a multi-layer Transformer network structure is adopted, with each layer containing a self-attention mechanism and a feed-forward network. Each layer of the Transformer consists of multi-head attention and point-wise feed-forward networks. The model configuration is as follows: number of transformer layers, 6; number of heads, 8; dimension of hidden layers, 512; dimension of feed-forward networks, 2048; dropout rate, 0.1. Such configuration ensures the model has sufficient capacity to process and learn complex text and knowledge graph data.

3.4.3. Graph Attention Mechanism

In this study, the graph attention mechanism is introduced, playing a pivotal role in the elaeagnus angustifolia disease detection smart agriculture system. This mechanism, different from the self-attention mechanism in traditional Transformers, is specifically designed for graph-structured data to enhance the model's capability to recognize and classify features of elaeagnus angustifolia diseases as shown in Figure 6.
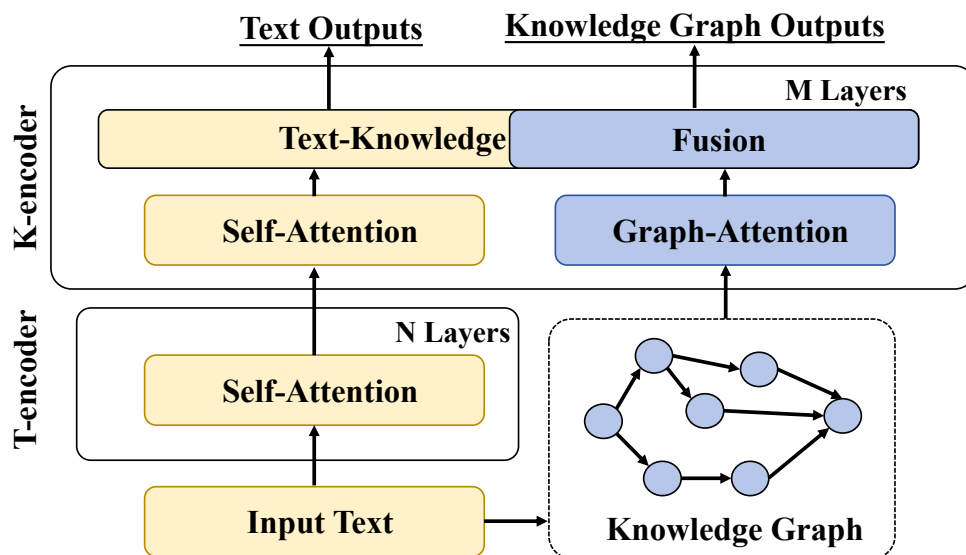

Figure 6. This figure illustrates the graph attention mechanism used in the smart agriculture system for elaeagnus angustifolia disease detection. It shows how traditional self-attention mechanisms are combined with graph attention mechanisms to process textual and knowledge graph data.

In traditional Transformer models, the self-attention mechanism focuses on the relationships between elements within sequence data, extracting features by calculating the weight distribution among elements in the sequence. This mechanism is suitable for processing linear data such as text; however, it cannot be directly applied to non-Euclidean

structures like graphs, where node connections are not merely linear or sequential. Graph attention is designed to address this issue, allowing the model to apply attention mechanisms on graph-structured data by focusing on complex relationships between nodes to optimize information extraction. Specifically, graph attention can dynamically compute attention weights for each node based on the features of its neighboring nodes, thereby capturing the relationships between nodes and the characteristics of the nodes themselves. The basic principle of a graph attention network (GAT) is to explicitly weight the features of adjacent nodes using the attention mechanism. The new feature representation for each node is the weighted sum of its neighbors' features, with weights dynamically determined by the attention mechanism. The mathematical formula is as follows:

$$\vec{h}_i' = \sigma\left(\sum_{j\in\mathcal{N}(i)} \alpha_{ij}\mathbf{W}\vec{h}_j\right) \tag{17}$$

where $\vec{h}_i'$ is the updated feature vector of node i, $\mathcal{N}(i)$ is the set of neighboring nodes of i, α_{ij} is the attention coefficient between node i and node j, $\mathbf{W}$ is a learnable weight matrix, and σ is a nonlinear activation function such as ReLU. The calculation of the attention coefficient α_{ij} typically involves a small neural network, which aims to assess the importance of the features of node j for node i:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\mathbf{a}^T[\mathbf{W}\vec{h}_i\|\mathbf{W}\vec{h}_j]))}{\sum_{k\in\mathcal{N}(i)}\exp(\text{LeakyReLU}(\mathbf{a}^T[\mathbf{W}\vec{h}_i\|\mathbf{W}\vec{h}_k]))} \tag{18}$$

The advantage of adopting the graph attention mechanism lies in its flexibility to handle complex relationships between nodes in graph-structured data, a capability that is challenging to achieve with traditional self-attention mechanisms. In the application of elaeagnus angustifolia disease detection, this means the system can dynamically adjust its focus on different disease markers based on specific features and contextual conditions of the diseases, thereby enhancing the accuracy and efficiency of diagnostics.

3.4.4. Graph Loss Function

In the elaeagnus angustifolia disease detection and smart agriculture system, an innovative graph loss function is designed to optimize the model's performance and fully utilize the structured information in the Agricultural Knowledge Graph. This loss function aims to improve learning efficiency and accuracy in processing graph-structured data, especially in tasks such as disease detection and classification. In traditional Transformer models, the loss function is usually a cross-entropy loss, which calculates the discrepancy between model outputs and true labels. However, this type of loss function does not directly consider the graph structural features of the data, such as the connectivity between nodes and edge information, which are particularly important when handling structured data like KGs. Therefore, the graph loss function designed not only calculates the label prediction loss for nodes but also includes a loss for the relationships between nodes to reinforce the model's learning of graph structures. Specifically, the graph loss function consists of two parts: node loss and edge loss. Node loss still utilizes cross-entropy loss, while edge loss is calculated by considering the similarity or relationship between pairs of nodes. The mathematical expression is as follows:

$$\mathcal{L}_{\text{graph}} = \mathcal{L}_{\text{node}} + \lambda\mathcal{L}_{\text{edge}} \tag{19}$$

where $\mathcal{L}_{\text{node}}$ is the node loss, calculated as:

$$\mathcal{L}_{\text{node}} = -\sum_{i=1}^{N} y_i \log \hat{y}_i \tag{20}$$

y_i is the true label of node i, and $\hat{y}_i$ is the model's predicted output for node i. $\mathcal{L}_{\text{edge}}$ is the edge loss, calculated as:

$$\mathcal{L}_{\text{edge}} = \sum_{(i,j)\in\mathcal{E}} \left\| \hat{y}_i - \hat{y}_j \right\|^2 \tag{21}$$

Here, $\mathcal{E}$ represents the set of all edges in the graph, and (i, j) is an edge connecting nodes i and j, with $\|\cdot\|$ representing the Euclidean distance. λ is a hyperparameter that balances the importance of the two parts of the loss. The design of the graph loss function is based on the assumption that connected nodes should have similar outputs, reflecting the similarity in relationships or functions between nodes in the graph. By minimizing edge loss, the model is encouraged to learn and maintain these intrinsic connections between nodes in the graph, thus better understanding and utilizing the information in graph structures. Mathematically, by introducing edge loss, an optimization problem is effectively addressed where the regularization term ensures that not only the predictive performance of individual nodes is considered but the structural characteristics of the entire graph are also maintained. This design effectively integrates global information of the graph structure into the model training process, helping to enhance the model's generalization ability on structured data. Applying the graph loss function to the task of elaeagnus angustifolia disease detection brings several advantages:

1. Enhanced Model Interpretability: By reinforcing the model's learning of relationships between nodes in the graph, the decision-making process of the model can be more easily interpreted, for instance, identifying which disease features are interconnected, providing valuable insights for agricultural experts in disease management and decision-making.
2. Improved Accuracy: As the graph loss function considers the relationships between nodes, the model can more accurately identify and classify nodes in the graph, which is particularly important in disease detection, potentially reducing cases of misdiagnosis and missed diagnosis.
3. Optimized Model Generalization: Considering the similarity between nodes during training helps the model to make accurate predictions even when faced with new types of diseases it has not previously encountered.

In summary, the graph loss function provides a powerful tool for the elaeagnus angustifolia disease detection and smart agriculture system, enabling the system not only to effectively handle graph-structured data but also to exhibit higher performance and better user experience in practical applications.

3.5. Evaluation Metrics

In this study, a variety of evaluation metrics are employed to comprehensively assess the performance of the elaeagnus angustifolia disease detection model and the smart agriculture system. These metrics, which include Precision, Recall, Accuracy, Mean Average Precision (mAP), and Frames Per Second (FPS), reflect the model's identification capability, classification efficiency, and operational efficiency from different perspectives and serve as essential tools for evaluating the model's overall performance. Precision measures the proportion of elaeagnus angustifolia disease images correctly identified as diseased by the model, i.e., the ratio of actual disease cases among all cases judged as diseased by the model. Recall, also known as the true positive rate, measures the proportion of actual disease cases captured by the model, i.e., the ratio of cases correctly identified by the model among all actual disease cases. Accuracy, the most intuitive performance metric, indicates the overall proportion of correct predictions made by the model, including both positive and negative classes.

The mAP is calculated by integrating the precision–recall curve, where precision at various recall levels is assessed to provide a comprehensive measure of model performance across different thresholds. The formula for mAP is as follows:

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^{n} \text{AP}_i \tag{22}$$

where n is the number of different thresholds, and AP_i is the average precision at the i-th threshold. This involves calculating the area under the precision–recall curve for each threshold, often achieved through numerical integration methods, then averaging these values. A higher mAP value indicates a more stable performance of the model across various operational points.

FPS, which refers to the number of frames transmitted per second, serves as a metric for evaluating the model's processing speed in practical applications. In smart agriculture applications, particularly in real-time monitoring and processing systems, a high FPS ensures the system's real-time responsiveness. The calculation formula for FPS is given by:

$$\text{FPS} = \frac{1}{T} \tag{23}$$

where T is the time required to process a single frame image. Through comprehensive analysis of these evaluation metrics, a thorough understanding of the model's performance in detecting elaeagnus angustifolia diseases is obtained, allowing for optimization and adjustment according to specific application needs. These metrics not only assist in identifying the strengths and weaknesses of the model but also guide future research on how to improve model architecture and training processes to achieve higher accuracy and efficiency.

3.6. Experimental Setup

3.6.1. Baseline

To evaluate the performance of the developed elaeagnus angustifolia disease detection model and smart agriculture system, several detection models are selected as baselines for comparison. These models include YOLOv5 [42], YOLOv8 [43], YOLOv9 [44], TinySegformer [12], RetinaDet [45], and EfficientDet [46]. By comparing these advanced baseline models with the developed system, not only is the effectiveness and advantage of our model in elaeagnus angustifolia disease detection validated but a deeper understanding of the performance differences among various models in handling complex agricultural images is also gained. This comparison not only aids in evaluating model performance but also guides future efforts in model optimization and algorithm adjustment.

3.6.2. Training Configuration

To ensure the reliability and generalizability of the evaluation results for the elaeagnus angustifolia disease detection model and smart agriculture system, both five-fold and ten-fold cross-validation methods were employed to train and test the model. Specifically, the dataset was divided into five non-overlapping subsets, with four subsets used for training and the remaining one used for testing in each iteration. This process was repeated five times, with each subset serving as test data once. Furthermore, to enhance the accuracy of the evaluation and the generalization testing of the model, ten-fold cross-validation was also conducted. In the ten-fold cross-validation, the dataset was divided into ten subsets, with nine subsets used for training and one subset used for testing in each iteration, repeated ten times.

Finally, to obtain more robust results, the outcomes of the five-fold and ten-fold cross-validation were averaged. This not only increased the reliability of the evaluation but also tested the stability and effectiveness of the model through multiple validation methods, ensuring the wide applicability and scientific validity of the evaluation results.

Concerning the training strategy, Stochastic Gradient Descent (SGD) was adopted as the optimization algorithm, as SGD has been demonstrated to achieve good convergence speed and optimization effects in many deep learning tasks. To prevent overfitting, dropout and L2 regularization techniques were introduced. The dropout rate was set at 0.5, and the L2 regularization coefficient was set at 1×10^{-4}. Additionally, to address potential issues of gradient vanishing or explosion during training, Batch Normalization was incorporated.

In this study, the settings for hyperparameters were determined through a series of preliminary experiments aimed at ensuring optimal stability and convergence during the model training process. Initially, the learning rate was set at 0.01. This value was chosen based on trials with various learning rates, which demonstrated that it effectively balanced the learning speed and model performance. To maintain model stability and convergence in the later stages of training, a learning rate decay strategy was implemented, whereby the learning rate was halved every 10 training epochs. This approach reduces the learning step size as the model approaches the optimal solution, thereby minimizing fluctuations near this solution.

Regarding the batch size, it was set to 32. This decision took into account the current GPU memory capacity and the complexity of the model. Larger batch sizes utilize more hardware parallelism, accelerating the training process, while smaller batches facilitate more precise adjustment of weights, preventing overfitting. The batch size of 32 was chosen based on a balance of model performance on the training and validation sets and computational resource constraints.

Furthermore, to monitor model performance and make necessary adjustments, an early stopping mechanism was introduced. Specifically, training automatically ceased if no significant improvement was observed on the validation set over 10 consecutive training epochs. This mechanism helps prevent overtraining of the model and conserves computational resources and time. These settings and optimizations enabled the model to exhibit good performance and stability under various training conditions.

3.6.3. Hardware Platform

In terms of hardware configuration, given the computationally intensive nature of training deep learning models, a computing platform equipped with high-performance GPUs was chosen. The specific setup included NVIDIA Tesla V100 GPUs (NVIDIA, Santa Clara, CA, USA), each with 32 GB of memory, capable of effectively supporting the training needs of large datasets and complex models. Additionally, the server was equipped with 128 GB of RAM and multi-core CPUs to ensure the efficiency of data processing and model training. Regarding software configuration, all experiments were conducted on a Linux operating system. Python was selected as the primary programming language, utilizing its robust library support for data processing and model building. Specifically, PyTorch 1.8 was used as the deep learning framework due to its flexible design interface and strong GPU acceleration capabilities. For data processing and enhancement, libraries such as NumPy 1.26, Pandas 1.5.3, and OpenCV 3.4.11 were also employed.

4. Results and Discussion
4.1. Disease Detection Results

The primary objective of this study was to evaluate and compare the performance of different models in the task of elaeagnus angustifolia disease detection, with the aim of assessing their practical applicability. The experimental results included five key metrics, Precision, Recall, Accuracy, mAP, and FPS, each reflecting the model's performance in disease detection. The table lists seven different models ranging from YOLOv5 to the method proposed in this article, each displaying performance data in detecting elaeagnus angustifolia diseases.

From Table 2, it is evident that the performance of the models generally shows a progressive improvement. Starting with YOLOv5, it achieved a precision of 0.82, recall of 0.80, accuracy of 0.81, mAP of 0.81, and an FPS of 21. This indicates that while YOLOv5 has

a solid foundation in speed and accuracy, it may struggle in more complex or demanding tasks. EfficientDet showed a slight improvement in performance, with a precision of 0.84, recall of 0.82, accuracy of 0.83, mAP of 0.83, and an FPS of 33, optimizing network structure and computational efficiency while maintaining high accuracy and improving processing speed. YOLOv8 and YOLOv9 further enhanced performance, with YOLOv8 achieving a precision of 0.86, recall of 0.84, accuracy of 0.85, mAP of 0.85, and an FPS of 28; YOLOv9 improved precision to 0.88, recall to 0.86, accuracy to 0.87, mAP to 0.87, and an FPS of 37. These models, inheriting the YOLO series' efficiency, incorporate more advanced network designs and optimization algorithms, facilitating better learning of elaeagnus angustifolia disease characteristics. TinySegformer and RetinaDet demonstrated even higher performance, particularly in complex environments. TinySegformer reached a precision of 0.90, recall of 0.88, accuracy of 0.89, mAP of 0.89, and an FPS of 36, while RetinaDet exhibited the highest precision of 0.92, recall of 0.90, accuracy of 0.91, mAP of 0.91, and an FPS of 45. These models, utilizing more complex network architectures and advanced image processing technologies, effectively enhance the detail recognition capability in disease detection.

Table 2. Disease detection results.

Model	Precision	Recall	Accuracy	mAP	FPS
YOLOv5 [8]	0.82	0.80	0.81	0.81	21
EfficientDet [7]	0.84	0.82	0.83	0.83	33
YOLOv8 [9]	0.86	0.84	0.85	0.85	28
YOLOv9 [10]	0.88	0.86	0.87	0.87	37
TinySegformer [12]	0.90	0.88	0.89	0.89	36
RetinaDet [11]	0.92	0.90	0.91	0.91	45
Proposed Method	0.94	0.92	0.93	0.93	57

The method proposed in this article showed the most exceptional results, with a precision of 0.94, recall of 0.92, accuracy of 0.93, mAP of 0.93, and an FPS of 57. This approach, integrating various advanced technologies and algorithm optimizations, achieves high accuracy while significantly accelerating processing speeds. From a mathematical perspective, this method likely employs more refined feature extraction techniques and an optimized attention mechanism, which aid the model in rapidly and accurately identifying and classifying diseases in practical applications. Theoretical analysis suggests that the improvements in model performance are primarily due to optimizations in deep learning architecture, computational efficiency, and a deepened understanding of elaeagnus angustifolia disease characteristics. The high FPS indicates the model's rapid response capability in practical applications, while high accuracy and recall ensure the model's reliability and practicality in actual agricultural environments. Each model's design and optimization aim to find the optimal balance between speed and accuracy in real-time disease detection, ensuring efficient and accurate support.

4.2. Detection Results Analysis

The aim of this experiment was to thoroughly analyze and compare the performance of various models in detecting different types of jujube diseases, thereby evaluating each model's accuracy in identifying specific diseases. This analysis not only revealed the strengths and limitations of each model in handling various diseases but also provided direction for future model optimization. The experiment covered common pests including the longhorn beetle, scale insect, psyllid, black leaf beetle, and mealybug—key threats to the healthy growth of jujube plants, as shown in Table 3.

Table 3. Detection results analysis.

Model	Longhorn Beetle	Scale Insect	Psyllid	Black Leaf Beetle	Mealybug
YOLOv5	0.81	0.79	0.81	0.83	0.81
EfficientDet	0.85	0.80	0.84	0.84	0.82
YOLOv8	0.88	0.82	0.86	0.85	0.84
YOLOv9	0.90	0.87	0.85	0.86	0.87
TinySegformer	0.87	0.88	0.89	0.91	0.90
RetinaDet	0.94	0.92	0.90	0.90	0.89
Proposed Method	0.94	0.93	0.92	0.92	0.94

According to the experimental results, there were significant differences in the performance of each model across various disease detection tasks. As seen in the table, the performance of the models generally improved with the novelty and complexity of the technologies used. As a baseline model, YOLOv5 achieved an accuracy of 0.81 for detecting longhorn beetles and displayed accuracies of 0.79, 0.81, 0.83, and 0.81 for scale insects, psyllids, black leaf beetles, and mealybugs, respectively. This shows that YOLOv5 has a solid baseline effectiveness in handling different diseases, though it may fall short in more complex or less frequent disease types. EfficientDet showed a slight improvement across all disease types, particularly with accuracies of 0.84 for both psyllids and black leaf beetles, indicating its capability in slightly more complex disease scenarios. YOLOv8 and YOLOv9 further improved accuracy, especially YOLOv9, which reached high accuracies of 0.87 for both scale insects and mealybugs, likely due to its deeper networks and more complex feature extraction capabilities. TinySegformer and RetinaDet exhibited even higher performance, particularly in recognizing mealybugs, with accuracies of 0.90 and 0.89, respectively. Their superior performance likely stems from their expertise in image segmentation, which is crucial in situations with rich detail or subtle manifestations of diseases. The proposed method achieved the highest accuracies across all disease detection tasks, especially with accuracies of 0.93 and 0.94 for scale insects and mealybugs, demonstrating the advantages of integrating multiple advanced techniques such as graph attention mechanisms and optimized loss functions.

Theoretically, the differences in performance between the models primarily originate from how they process images and learn features. Deep learning-based models like YOLO and EfficientDet rely on extensive data and deep networks to capture complex features, while TinySegformer and RetinaDet, through fine-grained image analysis techniques like segmentation networks, can handle detailed aspects of images more effectively. The proposed method enhances the capability to capture complex relationships between nodes (key features in images) by integrating graph attention mechanisms and optimized loss functions, thus achieving exceptional performance in various disease detection tasks. Moreover, the graph attention mechanism allows for the dynamic adjustment of attention allocation based on the relationships between nodes (features), which is crucial for identifying similar disease features or distinguishing diseases in complex backgrounds. This introduction of the mechanism enables the model not only to capture local features but also to integrate information at a global level, significantly enhancing the accuracy and efficiency of disease detection.

4.3. Attention Ablation Experiment Results

The primary aim of this experiment was to explore and compare the performance impact of different attention mechanisms in the task of elaeagnus angustifolia disease detection, to assess the specific contributions of each type of attention mechanism to model accuracy and efficiency. The experiment included three main types of attention mechanisms: self-attention, multi-head attention, and graph attention. The results are presented in Table 4.

Table 4. Attention ablation experiment results.

Model	Precision	Recall	Accuracy	mAP	FPS
Self-Attention	0.73	0.70	0.71	0.72	44
Multi-Head Attention	0.85	0.81	0.83	0.82	46
Graph Attention	0.94	0.92	0.93	0.93	57

The experimental results indicate significant differences in performance across the five metrics of precision, recall, accuracy, mAP, and FPS among the different attention mechanisms. Self-attention demonstrated relatively lower performance with a precision of 0.73, recall of 0.70, accuracy of 0.71, mAP of 0.72, and an FPS of 44. Self-attention mainly focuses on the relationships between elements within a single data stream, which may not adequately capture all critical information in complex graph-structured data, leading to poorer performance. The multi-head attention mechanism showed better performance, improving precision to 0.85, recall to 0.81, accuracy to 0.83, mAP to 0.82, and an FPS of 46. By processing information in parallel, multi-head attention can simultaneously focus on different aspects of the data across multiple subspaces. This design enables the model to understand data features more comprehensively, thereby enhancing performance. The graph attention mechanism exhibited the best performance in all tests, with a precision of 0.94, recall of 0.92, accuracy of 0.93, mAP of 0.93, and an FPS of 57. Designed specifically for processing graph-structured data, graph attention dynamically adjusts the weight of relationships between nodes, optimizing the understanding of node features and the overall graph structure. This mechanism is particularly suited for handling the complex relationships and non-linear features in agricultural disease detection, making the model more efficient and accurate in precisely identifying and classifying diseases.

Theoretically, while self-attention can handle dependencies in sequence data, its performance is limited in the face of graph-structured data due to its inability to directly map the complex spatial relationships between nodes. Multi-head attention, by extending the basic idea of self-attention, uses multiple independent attention "heads" to process data in parallel, enhancing the model's ability to capture information, especially in complex data with multiple associative properties. Graph attention further optimizes for the characteristics of graph data by calculating dynamic weights between nodes, which directly reflect the characteristics of the graph structure. Mathematically, this enhancement in the model's ability to encode the graph's topological structure significantly improves performance on structured data. Analyzing from the mathematical features of the model, the dynamic weight calculation formula in graph attention can more accurately reflect the actual strength of relationships between entities, which is crucial for enhancing the accuracy of disease detection. Moreover, the design of the graph attention network allows the model not only to learn local features of entities but also to integrate information across the entire graph through global optimization. This global perspective is unmatched by other attention mechanisms. Thus, through its mathematical and structural advantages, graph attention demonstrates superior performance in complex application scenarios like elaeagnus angustifolia disease detection.

4.4. Loss Function Ablation Experiment Results

The purpose of this experiment was to assess the impact of different loss functions on the performance of models in detecting elaeagnus angustifolia diseases, aiming to identify which loss function can more effectively enhance model performance in practical applications. The experiment compared three main types of loss functions, cross-entropy loss, focal loss, and graph attention loss, to reveal the effects of different loss functions on handling imbalanced datasets, enhancing model sensitivity to difficult-to-detect samples, and improving overall model performance. The results are displayed in Table 5.

Table 5. Loss function ablation experiment results.

Model	Precision	Recall	Accuracy	mAP	FPS
Cross-Entropy Loss	0.70	0.66	0.69	0.68	39
Focal Loss	0.85	0.81	0.84	0.84	47
Graph Attention	0.94	0.92	0.93	0.93	57

The experimental results demonstrate significant differences in performance across five key metrics—precision, recall, accuracy, mAP, and FPS—among the different loss functions. As a traditional loss function, cross-entropy loss showed relatively weaker performance in the task of elaeagnus angustifolia disease detection, with a precision of 0.70, recall of 0.66, accuracy of 0.69, mAP of 0.68, and an FPS of 39. Cross-entropy loss primarily focuses on improving the overall accuracy in classification tasks but fails to effectively address issues of data imbalance and the model's sensitivity to minority class samples. The focal loss function exhibited better results in this experiment, with precision increasing to 0.85, recall to 0.81, accuracy to 0.84, mAP to 0.84, and an FPS of 47. Designed to address class imbalance in classification tasks, focal loss adjusts the focus parameter within the loss function to make the model pay more attention to samples that are difficult to classify correctly, thereby enhancing the model's ability to recognize these samples. This is particularly important in practical applications like elaeagnus angustifolia disease detection, where disease manifestations vary widely, and some disease types may have fewer samples. The graph attention loss mechanism performed best in all tests, with a precision of 0.94, recall of 0.92, accuracy of 0.93, mAP of 0.93, and an FPS of 57. This loss mechanism integrates the characteristics of graph-structured data, enhancing the model's learning of complex relationships between nodes in the graph, making loss calculations more aligned with practical application needs. Especially in handling graph data, this loss mechanism effectively strengthens the model's capture of linkage information, significantly improving accuracy when identifying linkage-dense disease features.

Theoretically, while cross-entropy loss is widely used in multi-class classification problems, its limitations in specific application scenarios, particularly under imbalanced sample distribution, are evident. Focal loss enhances the model's predictive performance on minority classes by changing the weight distribution for difficult samples in the loss function, achieved mathematically by adjusting cross-entropy to increase the penalty for misclassified samples. Graph attention loss fundamentally alters the way loss is calculated by considering the dependencies between nodes to optimize the loss function. This method directly corresponds to the structural properties of graph data and serves as an effective supplement to traditional loss functions. Mathematically, graph attention loss optimizes the overall network structure by dynamically adjusting the weights of relationships between nodes, a method that has proven its importance in enhancing the performance of complex data processing both theoretically and practically.

4.5. Future Work

This study has achieved significant results in the research of elaeagnus angustifolia disease detection and smart agriculture systems, especially in applying different attention mechanisms and loss functions to improve the accuracy of disease recognition. However, there are some limitations in the research that need further improvement and exploration in future work. First, although the experimental results show that the model proposed in this paper outperforms existing methods on multiple performance indicators, these experiments were primarily conducted on fixed datasets. The complexity of actual agricultural environments far exceeds experimental conditions, and factors such as changes in lighting and background interference may affect the model's generalizability. Current research has not fully simulated the various challenges in real agricultural environments. Future studies need to validate the practicality and stability of the model in more diverse and complex environments. Secondly, although this study has considered the characteristics of

graph-structured data when designing the graph attention mechanism, the computational resources and time cost remain a challenge when dealing with large-scale graph data. The computational complexity of graph attention networks is relatively high, especially when the number of nodes and edges increases, which may limit the scalability and efficiency of the model. Future research needs to explore more efficient graph processing algorithms or develop new graph sampling techniques to enhance the training and inference speed of the model on large-scale datasets. Lastly, future research should also consider the practicality and cost effectiveness of model deployment. How to translate research findings into practically usable products, especially in resource-limited agricultural production environments, is an important research direction. This includes optimizing the storage and computational demands of the model to suit edge computing devices and developing user-friendly interfaces that allow non-professional users to easily use the system for disease detection and management decision-making.

5. Conclusions

This study is set against the backdrop of the rapidly evolving field of smart agriculture, focusing specifically on the automatic detection of elaeagnus angustifolia diseases. With the advancement of agricultural technology, precision agriculture is becoming a key technological route to enhance crop yield and quality. In this context, the use of advanced image processing and machine learning technologies for the detection and prevention of crop diseases, particularly in widely cultivated elaeagnus angustifolias, is of significant importance for ensuring food security and sustainable agricultural development.

The core innovation of this paper lies in the development of an efficient elaeagnus angustifolia disease detection and smart agriculture system that integrates the latest deep learning models and attention mechanisms, optimizing the entire process from image capture to disease diagnosis. By employing LLM, agricultural KGs, GNN, representation learning, neural-symbolic reasoning, and few-shot learning techniques, this study not only enhances the accuracy of disease detection but also provides richer decision support for agricultural disease management. In the experimental section, a series of experiments were designed to validate the effectiveness of the proposed model. Firstly, in the comparative experiment of different attention mechanisms, the graph attention mechanism (graph attention) performed the best with an average precision mean (mAP) of 0.93 and an accuracy rate of 0.93, significantly better than traditional self-attention and multi-head attention mechanisms. This result demonstrates that the graph attention mechanism can more effectively capture complex relationships between nodes in handling complex graph-structured data, providing strong technical support for precise disease identification. Secondly, the ablation experiment of different loss functions further showcased the superiority of the method proposed in this paper. Compared to traditional cross-entropy loss and focal loss, the graph attention loss mechanism proposed in this paper showed significant improvements in all performance indicators, particularly in precision and recall, reaching high scores of 0.94 and 0.92, respectively, fully demonstrating the effectiveness of optimizing loss calculations on graph-structured data. The main contribution of this paper is the successful integration of deep learning technologies with the field of agricultural disease detection, especially in the application of elaeagnus angustifolia disease detection. The effectiveness of the proposed model and algorithms was verified through meticulously designed experiments, not only enhancing the accuracy of disease detection but also providing a scientific basis for the early prevention and treatment of agricultural diseases. Furthermore, the models and algorithms in this study also offer new ideas and methods for research in related fields, especially showing great potential in handling unstructured agricultural data.

In conclusion, by introducing advanced machine learning technologies and algorithms, this paper provides an effective technical solution for disease detection tasks in smart agriculture, significantly enhancing the automation and intelligence level of detection. Future research will continue to explore more efficient algorithms and models to further

enhance the system's practicality and accuracy, expanding its application scope in actual agricultural production.

References

1. Zhang, Y.; Yang, X.; Liu, Y.; Zhou, J.; Huang, Y.; Li, J.; Zhang, L.; Ma, Q. A time-series neural network for pig feeding behavior recognition and dangerous detection from videos. *Comput. Electron. Agric.* **2024**, *218*, 108710. [CrossRef]
2. Lin, X.; Wa, S.; Zhang, Y.; Ma, Q. A dilated segmentation network with the morphological correction method in farming area image Series. *Remote Sens.* **2022**, *14*, 1771. [CrossRef]
3. Yang, C. Remote sensing and precision agriculture technologies for crop disease detection and management with a practical application example. *Engineering* **2020**, *6*, 528–532. [CrossRef]
4. Zhang, Y.; Wa, S.; Zhang, L.; Lv, C. Automatic plant disease detection based on tranvolution detection network with GAN modules using leaf images. *Front. Plant Sci.* **2022**, *13*, 875693. [CrossRef] [PubMed]
5. López-López, M.; Calderón, R.; González-Dugo, V.; Zarco-Tejada, P.J.; Fereres, E. Early detection and quantification of almond red leaf blotch using high-resolution hyperspectral and thermal imagery. *Remote Sens.* **2016**, *8*, 276. [CrossRef]
6. Liang, P.S.; Slaughter, D.C.; Ortega-Beltran, A.; Michailides, T.J. Detection of fungal infection in almond kernels using near-infrared reflectance spectroscopy. *Biosyst. Eng.* **2015**, *137*, 64–72. [CrossRef]
7. Lakshmi, R.K.; Savarimuthu, N. Investigation on object detection models for plant disease detection framework. In Proceedings of the 2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA), Arad, Romania, 17–19 December 2021; pp. 214–218.
8. Wang, H.; Shang, S.; Wang, D.; He, X.; Feng, K.; Zhu, H. Plant disease detection and classification method based on the optimized lightweight YOLOv5 model. *Agriculture* **2022**, *12*, 931. [CrossRef]
9. Qadri, S.A.A.; Huang, N.F.; Wani, T.M.; Bhat, S.A. Plant Disease Detection and Segmentation using End-to-End YOLOv8: A Comprehensive Approach. In Proceedings of the 2023 IEEE 13th International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 25–26 August 2023; pp. 155–160.
10. Malik, M.E.; Mahmud, M.S. Enhanced Weed Detection Using YOLOv9 on Open-Source Datasets for Precise Weed Management. In Proceedings of the 2024 ASABE Annual International Meeting, Anaheim, CA, USA, 28–31 July 2024; p. 1.
11. Sankareshwaran, S.P.; Jayaraman, G.; Muthukumar, P.; Krishnan, A. Optimizing rice plant disease detection with crossover boosted artificial hummingbird algorithm based AX-RetinaNet. *Environ. Monit. Assess.* **2023**, *195*, 1070. [CrossRef]
12. Zhang, Y.; Lv, C. TinySegformer: A lightweight visual segmentation model for real-time agricultural pest detection. *Comput. Electron. Agric.* **2024**, *218*, 108740. [CrossRef]
13. Zhang, Y.; Wa, S.; Liu, Y.; Zhou, X.; Sun, P.; Ma, Q. High-accuracy detection of maize leaf diseases CNN based on multi-pathway activation function module. *Remote Sens.* **2021**, *13*, 4218. [CrossRef]
14. Yao, Y.; Duan, J.; Xu, K.; Cai, Y.; Sun, Z.; Zhang, Y. A survey on large language model (llm) security and privacy: The good, the bad, and the ugly. *High-Confid. Comput.* **2024**, p. 100211. [CrossRef]
15. Osinga, S.A.; Paudel, D.; Mouzakitis, S.A.; Athanasiadis, I.N. Big data in agriculture: Between opportunity and solution. *Agric. Syst.* **2022**, *195*, 103298. [CrossRef]
16. Gregor, H.F. *Industrialization of US Agriculture: An Interpretive Atlas*; Routledge: Anaheim, CA, USA, 2021.
17. Wang, J.; Sun, Q.; Li, X.; Gao, M. Boosting language models reasoning with chain-of-knowledge prompting. *arXiv* **2023**, arXiv:2306.06427.
18. Nagasubramanian, G.; Sakthivel, R.K.; Patan, R.; Sankayya, M.; Daneshmand, M.; Gandomi, A.H. Ensemble classification and IoT-based pattern recognition for crop disease monitoring system. *IEEE Internet Things J.* **2021**, *8*, 12847–12854. [CrossRef]
19. Peng, R.; Liu, K.; Yang, P.; Yuan, Z.; Li, S. Embedding-based retrieval with llm for effective agriculture information extracting from unstructured data. *arXiv* **2023**, arXiv:2308.03107.

20. Tao, Z.; Wei, Y.; Wang, X.; He, X.; Huang, X.; Chua, T.S. Mgat: Multimodal graph attention network for recommendation. *Inf. Process. Manag.* **2020**, *57*, 102277. [CrossRef]
21. Zhou, J.; Li, J.; Wang, C.; Wu, H.; Zhao, C.; Teng, G. Crop disease identification and interpretation method based on multimodal deep learning. *Comput. Electron. Agric.* **2021**, *189*, 106408. [CrossRef]
22. Zhou, H.; Hu, C.; Yuan, Y.; Cui, Y.; Jin, Y.; Chen, C.; Wu, H.; Yuan, D.; Jiang, L.; Wu, D.; et al. Large language model (llm) for telecommunications: A comprehensive survey on principles, key techniques, and opportunities. *arXiv* **2024**, arXiv:2405.10825.
23. Zhai, Z.; Martínez, J.F.; Beltran, V.; Martínez, N.L. Decision support systems for agriculture 4.0: Survey and challenges. *Comput. Electron. Agric.* **2020**, *170*, 105256. [CrossRef]
24. Yahya, M.; Breslin, J.G.; Ali, M.I. Semantic web and knowledge graphs for industry 4.0. *Applied Sciences* **2021**, *11*, 5110. [CrossRef]
25. Chen, X.; Jia, S.; Xiang, Y. A review: Knowledge reasoning over knowledge graph. *Expert Syst. Appl.* **2020**, *141*, 112948. [CrossRef]
26. Tiwari, S.; Al-Aswadi, F.N.; Gaurav, D. Recent trends in knowledge graphs: Theory and practice. *Soft Comput.* **2021**, *25*, 8337–8355. [CrossRef]
27. Saiz-Rubio, V.; Rovira-Más, F. From smart farming towards agriculture 5.0: A review on crop data management. *Agronomy* **2020**, *10*, 207. [CrossRef]
28. Cravero, A.; Sepúlveda, S. Use and adaptations of machine learning in big data—Applications in real cases in agriculture. *Electronics* **2021**, *10*, 552. [CrossRef]
29. Sharma, A.; Jain, A.; Gupta, P.; Chowdary, V. Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access* **2020**, *9*, 4843–4873. [CrossRef]
30. Bhat, S.A.; Huang, N.F. Big data and ai revolution in precision agriculture: Survey and challenges. *IEEE Access* **2021**, *9*, 110209–110222. [CrossRef]
31. Arnaud, E.; Laporte, M.A.; Kim, S.; Aubert, C.; Leonelli, S.; Miro, B.; Cooper, L.; Jaiswal, P.; Kruseman, G.; Shrestha, R.; et al. The ontologies community of practice: A CGIAR initiative for big data in agrifood systems. *Patterns* **2020**, *1*, 100105. [CrossRef] [PubMed]
32. Nam, D.; Macvean, A.; Hellendoorn, V.; Vasilescu, B.; Myers, B. Using an llm to help with code understanding. In Proceedings of the IEEE/ACM 46th International Conference on Software Engineering, Lisbon Portugal, 14–20 April 2024; pp. 1–13.
33. Gao, M.; Hu, X.; Ruan, J.; Pu, X.; Wan, X. Llm-based nlg evaluation: Current status and challenges. *arXiv* **2024**, arXiv:2402.01383.
34. Wu, S.; Fei, H.; Qu, L.; Ji, W.; Chua, T.S. Next-gpt: Any-to-any multimodal llm. *arXiv* **2023**, arXiv:2309.05519.
35. Zhao, C.J. Agricultural knowledge intelligent service technology: A review. *Smart Agric.* **2023**, *5*, 126–148.
36. Teixeira, A.C.; Marar, V.; Yazdanpanah, H.; Pezente, A.; Ghassemi, M. Enhancing Credit Risk Reports Generation using LLMs: An Integration of Bayesian Networks and Labeled Guide Prompting. In Proceedings of the Fourth ACM International Conference on AI in Finance, Brooklyn, NY, USA, 27–29 November 2023; pp. 340–348.
37. Kumar, S.S.; Khan, A.K.M.A.; Banday, I.A.; Gada, M.; Shanbhag, V.V. Overcoming LLM Challenges using RAG-Driven Precision in Coffee Leaf Disease Remediation. In Proceedings of the 2024 International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS), Bengaluru, India, 22–23 April 2024; pp. 1–6.
38. Yang, S.; Yuan, Z.; Li, S.; Peng, R.; Liu, K.; Yang, P. GPT-4 as Evaluator: Evaluating Large Language Models on Pest Management in Agriculture. *arXiv* **2024**, arXiv:2403.11858.
39. Tzachor, A.; Devare, M.; Richards, C.; Pypers, P.; Ghosh, A.; Koo, J.; Johal, S.; King, B. Large language models and agricultural extension services. *Nat. Food* **2023**, *4*, 941–948. [CrossRef] [PubMed]
40. Ting, W.; Na, W.; Yunpeng, C.; Juan, L. Agricultural technology knowledge intelligent question-answering system based on large language model. *Smart Agric.* **2023**, *5*, 105.
41. Shutske, J.M. Harnessing the Power of Large Language Models in Agricultural Safety & Health. *J. Agric. Saf. Health* **2023**, *29*, 205–224.
42. Kim, J.H.; Kim, N.; Park, Y.W.; Won, C.S. Object detection and classification based on YOLO-V5 with improved maritime dataset. *J. Mar. Sci. Eng.* **2022**, *10*, 377. [CrossRef]
43. Wang, G.; Chen, Y.; An, P.; Hong, H.; Hu, J.; Huang, T. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios. *Sensors* **2023**, *23*, 7190. [CrossRef]
44. Wang, C.Y.; Yeh, I.H.; Liao, H.Y.M. Yolov9: Learning what you want to learn using programmable gradient information. *arXiv* **2024**, arXiv:2402.13616.
45. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
46. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; pp. 6105–6114.