



Article Research on Walnut (Juglans regia L.) Classification Based on Convolutional Neural Networks and Landsat-8 Remote Sensing Imagery

Jingming Wu^{1,2,†}, Xu Li^{1,2,†}, Ziyan Shi^{1,2}, Senwei Li^{1,2}, Kaiyao Hou^{1,2} and Tiecheng Bai^{1,2,*}

- ¹ College of Information Engineering, Tarim University, Alar 843300, China; 10757222273@stumail.taru.edu.cn (J.W.); lixu2866830@gmail.com (X.L.); 10757213118@stumail.taru.edu.cn (Z.S.); 10757222277@stumail.taru.edu.cn (S.L.); 10757222302@stumail.taru.edu.cn (K.H.)
- ² Key Laboratory of Tarim Oasis Agriculture, Ministry of Education, Tarim University, Alar 843300, China
- * Correspondence: baitiecheng@taru.edu.cn
- These authors contributed equally to this work.

Abstract: The study explores the use of convolutional neural networks (CNNs) and satellite remote sensing imagery for walnut analysis in Ganquan Township, Alar City, Xinjiang. The recent growth of walnut cultivation in Xinjiang presents challenges for manual data collection, making satellite imagery and computer vision algorithms a practical solution. Landsat-8 satellite images from Google Earth Engine underwent preprocessing, and experiments were conducted to enhance the ResNet model, resulting in improved accuracy and efficiency. Experiments were conducted to evaluate multiple CNN models and traditional methods, and the best detection method was chosen through comparisons. A comparison was drawn between traditional algorithms and convolutional neural network algorithms based on metrics such as precision, recall, f1-score, accuracy, and total time. The results indicated that although traditional methods were more efficient compared to CNN, they exhibited lower accuracy. In the context of this research, prioritizing efficiency at the cost of accuracy was deemed undesirable. Among the traditional algorithms employed in this study, k-NN produced the most favorable outcomes, with precision, recall, f1-score, and accuracy reaching 75.78%, 92.43%, 83.28%, and 84.46%, respectively, although these values were relatively lower than those of the CNN algorithm models. Within the CNN models, the ResNet model demonstrated superior performance, yielding corresponding results of 92.47%, 94.29%, 93.37%, and 93.27%. The EfficientNetV2 model also displayed commendable results, with precision, recall, and f1-score achieving 96.35%, 91.44%, and 93.83%. Nevertheless, it is worth noting that the classification efficiency of EfficientNetV2 fell significantly short of that of ResNet. Consequently, in this study, the ResNet model proved to be relatively more effective. Once optimized, the most efficient CNN model closely rivals traditional algorithms in terms of time efficiency for generating results while significantly surpassing them in accuracy. Through our studies, we discovered that once optimized, the most efficient CNN model closely rivals traditional algorithms in terms of time efficiency for generating results while significantly surpassing them in accuracy. In this study, empirical evidence demonstrates that integrating CNN-based methods with satellite remote sensing technology can effectively enhance the statistical efficiency of agriculture and forestry sectors, thus leading to substantial reductions in operational costs. These findings lay a solid foundation for further research in this field and offer valuable insights for other agricultural and forestry-related studies.

Keywords: walnut; convolutional neural networks; remote sensing; classification; residual network (ResNet)



Citation: Wu, J.; Li, X.; Shi, Z.; Li, S.; Hou, K.; Bai, T. Research on Walnut (*Juglans regia* L.) Classification Based on Convolutional Neural Networks and Landsat-8 Remote Sensing Imagery. *Forests* **2024**, *15*, 165. https://doi.org/10.3390/f15010165

Academic Editor: Cate Macinnis-Ng

Received: 3 December 2023 Revised: 8 January 2024 Accepted: 11 January 2024 Published: 12 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

1. Introduction

As urbanization continues to accelerate, large-scale agriculture is experiencing rapid transformation. The costs of relying solely on manual labor for essential crop-related tasks have been steadily increasing [1]. The relentless advancement and widespread adoption of smart agriculture have led to the use of technologies like artificial intelligence, which can now perform tasks that were previously conducted only by humans [2,3].

Walnut (Juglans regia L.) is a deciduous tree that can reach heights of 20-25 m [4,5]. The distribution of walnut-growing regions, including the United States, China, France, and Turkey, is primarily influenced by climatic and soil conditions. Different varieties of walnuts exhibit varying classifications based on their geographical origins, with Juglans regia, also known as the common walnut or Persian walnut, being one of the most widely cultivated and consumed species, with its origin tracing back to Asia, particularly in regions like Iran and China. At the same time, each variety has unique growth requirements typically found in specific geographic regions [6,7]. Adequate and evenly distributed rainfall during the growing season is essential for walnuts. Furthermore, a cold winter climate is conducive to walnut dormancy, and regions experience dry climates throughout the year [8]. Walnuts typically prefer neutral to alkaline soils with a pH range of 6.0 to 7.5, which aligns with the pH levels of many Xinjiang soils, making them suitable for walnut cultivation [9,10]. Additionally, due to the extensive utilization of drip irrigation in Xinjiang, the region can ensure an abundant water supply for sectors such as economic crop cultivation [11,12]. Consequently, Xinjiang represents the largest walnut-producing region in China and stands as one of the world's largest walnut-producing areas [13,14].

Walnuts, also an important economic crop in China, are cultivated in Xinjiang, where their cultivation plays a significant role in the economic and agricultural development of the region [15]. Furthermore, the cultivation of walnuts contributes to soil conservation and water resource management, effectively combating desertification and promoting environmental protection [16,17]. This multifaceted impact underscores the vital role of walnut cultivation in not only economic development but also environmental sustainability in Xinjiang.

Research efforts in the fields of agriculture and forestry have increasingly turned towards the exploration of machine learning, cloud computing, and remote sensing, reflecting a significant level of enthusiasm [18]. Remote sensing imagery, a critical component of these efforts, holds extensive applications in environmental monitoring, agriculture, urban planning, geological exploration, and weather forecasting, among other fields. This comprehensive geographical information facilitates informed decision-making and analysis, proving particularly valuable for resource management and environmental protection. Typically stored in raster format, each pixel in remote sensing imagery provides specific geographic information, enabling diverse analyses such as image classification, change detection, and feature extraction [19,20]. An illustration of the significance of remote sensing technology in agriculture and forestry is found in several notable studies. Yifan Bo et al. investigated the integration of cloud computing and Internet of Things technologies into the agricultural and forestry sectors, highlighting the feasibility, applications, and prospects of this fusion [21]. In another study, Shanwen Zhang et al. utilized deep belief networks (DBNs) to develop a predictive model for winter jujube pest and disease forecasting, achieving an impressive 84% accuracy rate [22]. Weijia Li et al. achieved a remarkable 96% accuracy rate in oil palm tree detection and counting through the integration of deep learning with high-resolution remote sensing imagery [23]. Remote sensing, as a technology that acquires information about the Earth's surface from distant sensors like aircraft or satellites, has been a focal point of extensive research [24]. An illustration of this is the study by Nicholas F. McCarthy et al., who employed deep learning techniques to enhance the clarity of geostationary satellites, providing crucial support for decision-making in the context of high-impact wildfires [25]. These developments highlight the growing importance of remote sensing in driving advancements in agriculture and forestry, as well as its potential to revolutionize decision-making processes in these fields.

Google Earth Engine is a cloud platform developed by Google with the purpose of supporting large-scale geospatial data analysis and visualization. By integrating remote sensing imagery, geospatial data, and computing capabilities, it provides a vast collection of Earth observation data. Additionally, GEE supports the processing and analysis of multispectral and hyperspectral data, which makes it applicable to tasks such as remote sensing, vegetation monitoring, and land use change analysis. Moreover, interactive map visualization capabilities are offered by GEE, enabling immediate inspection and exploration of geospatial data, thereby enhancing users' ability to gain a deeper understanding and interpretation of analytical outcomes [26,27]. The system finds extensive applications and potential in the fields of geospatial data science, environmental research, and sustainable development [28–31]. It has notably addressed the issue of insufficient resolution in satellite imagery, especially in the proliferation of relevant research areas and large-scale agricultural models worldwide.

In recent years, the continuous advancements in artificial intelligence, cloud computing, and other technologies have enabled computers to gather a vast amount of information from remote sensing images, surpassing the direct capabilities of human eyes [32]. This progress in technology, coupled with the ongoing enhancement in image resolution, has led to a steady growth in the information extracted via computers. As a result, there is now a solid foundation for employing diverse methods to achieve more refined classification and recognition of remote sensing images [33]. Many researchers have conducted studies in this area by combining relevant algorithms with satellite imagery, showcasing the potential for further advancements in the field [34].

Our previous research focused on the extraction of walnut areas and growth monitoring through the integration of satellite remote sensing imagery, vegetation indices, and machine learning techniques, which yielded favorable results [35]. While our earlier study heavily relied on band data and time series analysis, the current research places primary emphasis on remote sensing imagery, feature extraction, and the associated algorithms. In this study, we combined satellite remote sensing imagery with convolutional neural network algorithms. We began by obtaining relevant satellite remote sensing images through the necessary operations on the Google Earth Engine (GEE) platform. Subsequently, these images were processed further to prepare for the application of relevant algorithms in the classification of walnut images. During the study, we employed convolutional neural network models, including AlexNet, Visual Geometry Group (VGG), GoogleNet, ResNet, and EfficientNetV2, for image classification. Moreover, we utilized traditional methods, such as BP neural network, k-NN, PCA, LDA, and SVM, to classify walnuts and compared the results with those obtained using convolutional neural networks. Ultimately, this study was able to identify the relatively superior classification methods.

The application of convolutional neural networks (CNNs) in conjunction with satellite remote sensing imagery for walnut classification demonstrates significant agricultural potential, given the advancing technologies and the increasing prevalence of large-scale farming. The approach harnesses remote sensing and deep learning technologies for walnut classification, enhancing land and resource management and advancing the automation of agricultural practices, thereby promoting sustainable agriculture. The Landsat-8 satellite provides multispectral remote sensing imagery, and convolutional neural network (CNN) models effectively capture intricate vegetation features, thereby improving classification accuracy and furnishing valuable insights into land cover and vegetation health. This methodology is not limited to walnut classification but can be readily applied to the classification and monitoring of a wide range of crops, underscoring its broad practical utility.

The experiment aimed to classify walnut plants in Landsat-8 remote sensing images to automate monitoring and classification. This can relieve the burden of manual classification, increase efficiency, and potentially positively impact agricultural management in large-scale walnut production areas. Additionally, it allows for precise field management, providing a scientific basis for agricultural production, optimizing agricultural resource utilization, and improving yield and quality, thus supporting environmental protection and sustainable agriculture. The second part describes the research area and methods used; the third part presents the experimental results; the fourth part discusses the research findings, further analyzing the experiment's strengths, limitations, and future research directions; and the fifth part reviews the research results, explaining the research significance and future work.

2. Materials and Methods

2.1. Research Area and Image Acquisition

2.1.1. Research Area

Ganquan Township, situated in Awati County, Aksu Region, Xinjiang, is under the jurisdiction of the Nongyi Division of Alar City (latitude $40^{\circ}22'30''$ N, longitude $80^{\circ}03'45''$ E) on the northwest edge of the Taklamakan Desert. It is governed by the First Agricultural Bureau of Aral City. The region experiences arid, temperate, continental, and desert climates, which are characterized by abundant sunshine, significant thermal resources, and large diurnal temperature variations. The annual average temperature is 11.53 °C, with a maximum temperature of 43.9 °C and a minimum temperature of -28.8 °C. The average frost-free period lasts for 195 days, with an annual accumulated temperature of 4620.8 °C. Furthermore, the annual average solar radiation is 142 kcal/square centimeter, and the annual average wind speed is 21 m per second. This area also receives an annual average of 2793.4 h of sunshine, 73.5 mm of precipitation, and 1748.75 mm of evaporation. Figure 1 depicts the geographical characteristics of the area.



Figure 1. Geographic map of the study area.

2.1.2. Data Information and Experimental Environment

The research area was meticulously surveyed on-site in the pursuit of data collection, with the employment of Global Positioning System (GPS) equipment for precise geolocation. Relevant data from remote sensing imagery was used to subdivide the entire agricultural and forested landscape of the research area into a grid of plots. Data points were then collected at random locations within each grid, utilizing GPS devices. Field sampling commenced in June to account for the unique nature of crops within the agricultural and forest categories, ensuring the consistency of the cultivated crops throughout the study period and guaranteeing the accuracy of the sampling process.

The Landsat-8 Level-2 data, with a 30 m resolution, covering different areas of Ganquan Town, Alar City, in the Xinjiang Uygur Autonomous Region, were used within this study. Originally, consideration was given to incorporating Sentinel-2 satellite data into the research. However, issues emerged during the acquisition and preprocessing of the Sentinel-2 imagery, leading to missing image data and associated spectral information within the study area. Despite attempts to obtain and process relevant imagery from the European Space Agency and the Google Earth Engine (GEE) platform for experimentation, the final results varied significantly when compared to Landsat-8 data. As a result, it was decided to continue the experiments using Landsat-8 imagery. The study area's intricate nature, comprising encompassing diverse environmental imagery, led to initial experiments using only a subset of spectral bands, resulting in subpar classification results. Consequently, seven spectral band images (B2, B3, B4, B5, B6, B7, and B9) were collected and utilized for subsequent classification research. Given the presence of various disturbances in the study area, such as rivers, canals, buildings, deserts, and diverse vegetation, individual spectral band combinations yielded unsatisfactory results, prompting the selection of multiple spectral band images for the study. Ultimately, it was discovered that four specific bands, B2, B5, B6, and B7, exhibited higher sensitivity, significantly enhancing detection accuracy when combined with the B3 band. Furthermore, the inclusion of other bands also contributed to improvements, albeit to a relatively lesser extent. The selected imagery from August to September 2022 corresponds to the rapid lipid conversion phase and fruit maturation period of walnuts [36]. Precise delineation of all walnut areas in Ganquan Town and non-walnut areas within the research region was conducted, and corresponding vectors were drawn and integrated using the combination bands of Landsat-8 satellites in Google Earth. The selected band images were merged in the study area using Google Earth Engine. Cloud removal and correction operations were subsequently carried out to enhance the image quality. The resultant images were partitioned into a grid of 256×256 pixels, each with a 30 m resolution, for further analysis [37]. All data related to walnut crops, other vegetation, and architectural water flows within the study areas were meticulously collected through manual field surveys and mapped accordingly. Following calibration and adjustment, satellite remote sensing images from different spectral bands were merged to generate corresponding images. Ultimately, a total of 3369 experimental images were obtained from August to September 2022, comprising 1352 images depicting walnut orchards and 2017 images representing interfering factors. These images were meticulously categorized into distinct training, validation, and test sets, serving the purpose of the research. Owing to the intrinsic complexities found within the imagery, such as regions entirely bereft of walnut trees or densely populated with them, as well as disruptive elements like cloud cover, coupled with the varying tile and row values across different latitudinal and longitudinal study zones, the acquired image count exhibited fluctuations. The tile count ranged from 4 to 16, while the row count spans from 13 to 24. Due to data constraints, the image classification ratio was set at 6:1:3.

2.1.3. Methodological Models Used in the Study

The research in this study involved using convolutional neural network algorithms in conjunction with satellite remote-sensing imagery. Initially, the imagery underwent radiometric calibration, atmospheric correction, cloud removal, and other preprocessing tasks using the Google Earth Engine (GEE). Subsequently, the annotated images were generated based on ground-truth information obtained in the field. Some of these labeled images are presented in Appendix A, Figure A1.

Following image preprocessing, various models, including AlexNet, Visual Geometry Group (VGG), GoogleNet, ResNet, and EfficientNetV2, were developed for image classification. In addition to CNN models, traditional methods such as BP neural networks, k-NN, PCA, LDA, and SVM were incorporated for walnut classification. The performance metrics of these methods were then compared to identify the most optimal approach.

2.2. Convolution Neural Network

One of the representative algorithms of deep learning, a convolutional neural network (CNN), is a type of feedforward neural network with a deep structure that incorporates convolution calculations [38,39]. The CNN model is designed with specific layers: an input layer, a convolution layer, a pooling layer, a fully connected layer, and an output layer.

This study utilizes various CNN models, enhancing and refining them to maximize the classification accuracy of the relevant data.

2.2.1. AlexNet

AlexNet, designed by Alex Krizhevsky et al. [40], has been a pioneering model in the development of deep learning, particularly in the field of computer vision. This deep convolutional neural network (CNN) has profoundly influenced subsequent deep learning models, with its design principles and architecture demonstrating the effectiveness of deep CNNs and laying the foundation for subsequent research [41]. The original code for the program can be accessed at https://paperswithcode.com/method/alexnet (accessed on 1 March 2023). The AlexNet model made an early appearance in the landscape of convolutional models, but it marked a significant milestone as the first model to successfully incorporate techniques such as ReLU activation, Dropout, and LRN within the realm of CNNs. These innovations not only distinguished AlexNet but also served as a rich source of inspiration for numerous subsequent models seeking improvements. In our research, we optimized the model through adjustments, making it a valuable reference point for comparative analysis in this study. Significantly, compared to traditional shallow networks, AlexNet adopts a deeper architecture that allows the network to learn more complex and abstract feature representations. Additionally, AlexNet enhances performance by increasing the width of the network through the addition of more neurons. It introduces large-sized filters to capture local features in images and, for the first time, incorporates the rectified linear unit (ReLU) as the activation function, achieving notable results. However, the performance of the original model was unsatisfactory in this study, leading to adjustments in the model and resulting in the final model structure, as illustrated in Figure 2. The input image size used for the experiments is 256×256 , with a convolution kernel of 3 and a step size of 1. In the model, an initial ZeroPadding2D operation was applied to mitigate information loss, adding one row of zero pixels to the input image horizontally and vertically, along with one column of zero pixels on the right side and bottom. Subsequently, convolutional layers, pooling layers, a flattening layer, a Dropout layer, and dense layers were incorporated. In the initial convolution, 48 convolution kernels of size 11×11 and a stride of 4 were employed using the rectified linear unit (ReLU) activation function. Subsequent adjustments were made in the following pooling and convolution layers, predominantly using 3×3 and 5×5 convolution and pooling kernels. The activation function employed throughout was ReLU, with a dropout rate of 0.2 in the dropout layers. Finally, the output was obtained using the Softmax activation function, aiming to enhance the extraction of pertinent features and achieve superior classification performance. The algorithm was fine-tuned with a learning rate of 0.0005 and a batch size of 32, and performance stability was observed after approximately 280 training iterations.



Figure 2. AlexNet model.

The Visual Geometry Group (VGG) is a deep convolutional neural network (CNN) architecture proposed by the Visual Geometry Group [42]. It is characterized by its depth and the use of very small convolutional kernels for its convolutional operations, distinguishing it from previous network architectures. This design choice increases the network's depth and allows for the stacking of multiple small-sized convolutional kernels to capture richer image features. In addition, VGG incorporates max-pooling layers to reduce the spatial size of feature maps and extract more salient features [43]. The original code for the VGG program can be accessed at https://paperswithcode.com/method/vgg (accessed on 1 March 2023).

The experiment utilized the VGG19 architecture, where the feature extraction layers were created using the "feature" function, followed by the application of ReLU activation and a dropout rate of 0.4. The final output was obtained by applying the Softmax activation function. Within the feature extraction layers, 3×3 kernels with "SAME" padding were utilized, and ReLU activation was applied. The algorithm was fine-tuned with a learning rate of 0.0001 and trained with a batch size of 32. Performance stability was observed after approximately 220 training iterations.

2.2.3. GoogleNet

GoogleNet, a deep convolutional neural network (CNN) architecture proposed by the Google team [44,45], aims to tackle the challenges of parameter count and computational complexity in deep networks while also improving network performance.

GoogleNet introduced the Inception module, a multi-scale convolutional structure that processes input feature maps with different scale convolutional kernels in parallel and then concatenates the results. This approach enhances the network's representational capacity by enabling the learning of features at different levels. A key aspect of GoogleNet's architecture is the extensive use of 1×1 convolutional kernels to reduce the number of network parameters. By employing 1×1 convolutions, the network is able to efficiently reduce the feature map channels, resulting in a reduction of the computational load in subsequent layers. The availability of the original code for the program can be found at https://paperswithcode.com/method/googlenet (accessed on 1 March 2023).

The architecture of GoogleNet, including the design principles and the introduction of the Inception module and 1×1 convolutional layers, has significantly advanced subsequent deep learning models. This approach effectively constructs deeper and wider network structures while maintaining a balance between the number of parameters and computational complexity. The model used in this study is shown in Figure 3 and entails the construction of the GoogLeNet model. The experiment involves using relevant images as input, employing 3×3 convolutional kernels with ReLU activation, and defining Inception modules that include convolution and pooling operations. The InceptionAux class incorporates average pooling, convolution, batch normalization, ReLU activation, fully connected layers, and Softmax activation. For the experiment, a batch size of 32 is utilized, with the learning rate set to 0.0004. The model stabilizes after approximately 200 training epochs.



Figure 3. GoogleNet model.

2.2.4. ResNet

In 2015, Microsoft Research introduced ResNet, a deep convolutional neural network (CNN) architecture [45] designed to alleviate the problems of vanishing gradients and degradation in deep networks, thereby facilitating easier training and optimization of the network.

ResNet, a deep learning architecture, addresses the vanishing gradient problem in traditional convolutional networks by incorporating residual blocks with skip connections, also known as shortcut connections. These skip connections enable the direct addition of the input feature mapping to the output feature mapping within each residual block, thereby facilitating the easier propagation of information. Consequently, the network is capable of skipping certain layers when necessary, thus avoiding information loss and degradation. The use of residual blocks in ResNet allows the network to have increased depth, reaching dozens or even hundreds of layers without facing degradation issues, making it easier to extend the network's depth. Following the final residual block, the architecture typically employs global average pooling to convert the feature map into a vector, complemented by a fully connected layer for classification or regression tasks. This design not only ensures the network's simplicity but also reduces the number of parameters [46]. The original code for the implementation of ResNet can be found at https://paperswithcode.com/method/resnet (accessed on 1 March 2023).

The significant impact of the design principles and architecture of ResNet in addressing the issue of vanishing gradients and degradation in deep networks has been widely acknowledged. The success of ResNet has not only demonstrated the feasibility of constructing deeper and more powerful networks through the use of skip connections but has also established itself as a cornerstone for subsequent deep-learning models. The emergence of continuous improvements and variants of ResNet, such as ResNet-34, ResNet-50, and ResNet-101, has further solidified its standing, showcasing outstanding performance in various computer vision tasks. For this study, three variants of ResNet—ResNet-34, ResNet-50, and ResNet-101—were chosen for experimentation. The final model of ResNet utilized in this experiment is illustrated in Figure 4a, while the structure of the sequential model is depicted in Figure 4b. The experimental setup involved the implementation of the BasicBlock and Bottleneck classes with specific configurations. In the BasicBlock, adjustments were made to the convolutional kernel size to 3 with a stride of 1, along with the application of the ReLU activation function. Conversely, the Bottleneck class incorporated multiple convolutional layers with varying strides, applying batch normalization after each convolution to ensure normalized activations, which in turn stabilized training and improved convergence performance. Throughout the experiment, ReLU activation was consistently employed, and downsampling was implemented in subsequent layers to maintain a consistent channel and shape between input and output. The experiment also incorporated a dropout rate of 0.5, a batch size of 16, and a learning rate of 0.0002. Convergence was achieved after approximately 110 training iterations.



Figure 4. ResNet model. (a) ResNet model structure. (b) Sequential model structure.

2.2.5. EfficientNet

In 2019, the Google Brain team proposed EfficientNet, an efficient convolutional neural network architecture aimed at improving performance and computational efficiency by balancing network depth, width, and resolution. This architecture introduces a compound coefficient that scales the depth, width, and resolution of the network simultaneously, allowing for balanced expansion under different resource constraints. EfficientNet utilizes MBConv (Mobile Inverted Residual Bottleneck) blocks as its basic building units, which include depthwise separable convolutions, pointwise convolutions, and skip connections, resulting in highly efficient computation with fewer parameters and effective feature representation learning. The network structure comprises multiple repeated stages, each containing several MBConv blocks. As the network depth increases, the resolution of the feature maps gradually decreases while the number of channels increases. Consequently, this progressive design enables the network to effectively learn features from images of different resolutions. The original code for the program is available at https://paperswithcode.com/method/efficientnet (accessed on 1 March 2023).

The main focus of this study was to experiment with EfficientNet-V2, an improved version of EfficientNet. EfficientNet-V2 combines training-aware neural architecture search (NAS) and scaling techniques, resulting in a smaller size and faster speed compared to previous networks [47]. It also introduces progressive learning, which adaptively adjusts the regularization strength based on image size, accelerating training while enhancing accuracy. To conduct the experiments, EfficientNet models were created using the Keras library. These models were configured with various parameters, including zero-padding in 2D convolutions, the definition of inverted residual blocks, specified width and depth, and the use of separate functions for experimentation. A learning rate of 0.01 and a dropout rate of 0.2 were employed in the process. For the experimentation with EfficientNet-V2, adjustments were made to the convolutional kernels in the SE layer, MBConv layer, and stem layer. The primary kernel size utilized was 3, and the MBConv layer was expanded to improve accuracy. Additionally, the learning rate was set to 0.01, and a dropout rate of 0.2 was utilized. It is noteworthy that training stabilized after approximately 380 iterations.

2.3. Traditional Methods

In addition to the CNN algorithm, this study utilizes several traditional approaches for the classification of walnuts using satellite remote sensing. These traditional approaches include principal component analysis (PCA) [48], linear discriminant analysis (LDA) [48], backpropagation neural network (BP) [49], support vector machine (SVM) [50], and k-nearest neighbor classification (k-NN). Despite the excellent classification performance demonstrated by traditional methods, they have been incorporated alongside CNN in this study to achieve a comprehensive evaluation of classification approaches in satellite remote sensing for walnuts.

(1) The PCA algorithm is a method for feature extraction that utilizes orthogonal transformation to convert the observed data, initially represented by linearly correlated variables, into a reduced set of linearly independent variables known as principal components. The formula for PCA, as shown in Equation (1), involves several key variables: x_i represents the original sample point, z_{ij} represents the coordinate projection, w_i represents the standard orthogonal basis vector, z_i represents the projection of x_i in the low-dimensional coordinate system, z_{ij} represents the j-dimensional coordinate of x_i in the low-dimensional coordinate system, and const represents the constant value of $x_i x_i^T$.

$$z_{ij} = w_i^T x_i$$

$$\sum_{i=1}^{m} ||\sum_{j=1}^{d'} z_{ij} w_i - x_i|_2^2 = \sum_{i=1}^{m} z_i^T z_i - 2 \sum_{i=1}^{m} z_i^T w^T x_i + const$$

$$\to tr(w^T(\sum_{i=1}^{m} x_i x_i^T)w)$$
(1)

The LDA algorithm is a classical binary classification method that encompasses techniques such as the gamma function, binomial distribution, conjugate distribution, Gibbs sampling, and other statistical methods. It performs dimensionality reduction by mapping all samples to a one-dimensional coordinate axis, followed by the establishment of a threshold to distinguish the samples.

The binomial distribution is a discrete probability function that characterizes the number of successes in n independent experiments, each with a probability of *p*.

(2) The fundamental model of the SVM algorithm is a linear classifier with the largest margin defined in the feature space, making it a two-classification model. Its learning strategy is focused on maximizing this margin, and this objective can be mathematically formulated as solving convex quadratic programming. This is also equivalent to minimizing the regularized hinge loss function. The learning algorithm of SVM is thus centered around solving convex quadratic programming through optimization techniques. A crucial aspect of SVM lies in determining geometric spacing and addressing the issue of linear separability.

(3) This study utilizes the backpropagation (BP) and k-nearest neighbor (k-NN) classification algorithms. The BP algorithm encompasses two primary processes, namely, the forward propagation of signals and the backpropagation of errors. The error backpropagation algorithm has gained widespread usage in training multilayer feedforward networks, leading to the multilayer feedforward network being commonly referred to as the BP network.

The k-NN algorithm involves identifying the K instances closest to the new input instance in the training dataset and classifying the input instance into the majority class of the K instances.

2.4. Evaluation of Precision and Efficiency

In this study, precision evaluation metrics such as f1-score, precision, recall [51], and accuracy were utilized (Equations (2)–(5)) to assess the performance. These metrics are computed by taking into account the number of correctly detected targets (*TP*), incorrectly detected targets (*FP*), and targets that were not detected at all (*FN*), as represented in the equations.

$$F1 = (2 \times P \times R) / (P + R)$$
⁽²⁾

$$P = TP/(TP + FP) \tag{3}$$

$$R = TP/(TP + FN) \tag{4}$$

$$Acc = (TP + TN)/(TP + FP + FN + TN)$$
(5)

When selecting different algorithms, it is important to consider not only their performance in achieving optimal results, but also their detection efficiency and the number of training iterations required. This comparison should take into account the average time required per training, the number of training iterations, testing time, and the total time. Different algorithms may require varying numbers of training iterations to achieve optimal results, making it essential to include this consideration in the evaluation process.

3. Results

3.1. AlexNet Results

The AlexNet algorithm initially performed poorly due to issues with the original model's parameters and structure. Consequently, modifications were made to the model in an attempt to improve its performance. Despite tuning the corresponding parameters, the final accuracy remained unsatisfactory, prompting further adjustments to the model's architecture. To mitigate underfitting during classification, additional convolutional layers were incorporated, while pooling and dropout layers were introduced after the convolutional layers to address overfitting. Furthermore, dense layers were added to strengthen the connections between relevant feature values. As a result of these modifications, the

precision, recall, f1-score, and accuracy improved to 96.81%, 87.57%, 91.96%, and 91.58%, respectively. For detailed results, the improved AlexNet model's confusion matrix can be referred to in Appendix A, Figure A2, included herein.

3.2. VGG Results

In this study, it was observed that the original parameters of the VGG model yielded unsatisfactory performance. As a result, adjustments were made to the model's parameters using the experimental images. Subsequently, fine-tuning the parameters led to a substantial improvement in the classification results, as indicated by the precision, recall, f1-score, and accuracy values of 90.98%, 87.98%, 89.46%, and 89.41%, respectively. The specific outcomes of the enhanced VGG model's confusion matrix can be found in Appendix A, Figure A3.

3.3. GoogleNet Results

After conducting initial experiments using the original GoogleNet model, it was found that the model's performance was suboptimal. Consequently, adjustments were made to the model's structure by incorporating additional convolutional layers and two inception layers at the end of the original architecture. As a result of these modifications, a notable enhancement in the model's fitting performance was observed. However, it became apparent that overfitting was occurring post-adjustment, leading to a subsequent introduction of flattening, dense, dropout, and pooling layers to mitigate this issue. The subsequent experiments yielded promising outcomes, with the precision, recall, f1-score, and accuracy values achieving 87.10%, 95.23%, 90.98%, and 90.99%, respectively. A comprehensive breakdown of the enhanced GoogleNet model's confusion matrix is provided in Appendix A, Figure A4.

3.4. ResNet Results

The initial classification study of walnuts in the ResNet model yielded only moderate performance, prompting adjustments to the relevant model parameters based on experimental data. Subsequent experiments were conducted using ResNet34, ResNet50, and ResNet101 models, with the latter demonstrating the best performance, albeit still comparably inferior to the other models. To rectify this, modifications and adjustments were made to the sequential model, accompanied by the addition of an extra sequential structure. Following multiple experiments, the final results exhibited increased stability and improved performance. Specifically, the precision, recall, f1-score, and accuracy values were measured at 92.47%, 94.29%, 93.37%, and 93.27%, respectively. The detailed results of the enhanced ResNet model's confusion matrix are presented in Appendix A, Figure A5.

3.5. Efficientnet Results

In experimental trials using the EfficientNet model, it was evident that the use of EfficientNetV2 led to notably superior outcomes compared to the original EfficientNet model. The performance of the original model in the experiments was subpar, necessitating modifications to the model's architecture. To enhance classification outcomes, the model was enhanced by the addition of five Fuse-MBConv modules and thirty extra MBConv modules. As a result, the model achieved improved precision, recall, f1-score, and accuracy values of 96.35%, 91.44%, 93.83%, and 93.47%, respectively. For a detailed breakdown of the enhanced EfficientNet model's confusion matrix, please refer to Appendix A, Figure A6.

3.6. Convolutional Neural Network Classification Efficiency Summary

Table 1 presents the classification efficiency results for different convolutional neural network models. The data shows that the ResNet and AlexNet models achieved the highest efficiency, while the EfficientNetV2 model showed relatively lower efficiency.

Model	Average Length of Time	Number of	Test Duration (s)	Total Duration (s)
	Ter manning Session (3)	mannings	Duration (9)	Duration (9)
AlexNet	12	70	6	846
VGG	19	180	7	3427
GoogleNet	21	70	8	1478
ResNet	23	30	8	698
EfficientnetV2	10	460	11	4611

Table 1. Summary of classification efficiency.

3.7. Results of Traditional Classification Methods

In the experimentation with the backpropagation (BP) neural network, superior outcomes were achieved by first adjusting pertinent parameters such as the learning rate, activation function, and weights, followed by numerous iterations of experiments. Similarly, the k-nearest neighbors (k-NN) algorithm underwent adjustments primarily to the value of K, distance metric, and weighting, ultimately resulting in the selection of K = 5 and the utilization of the Euclidean distance metric, which produced the optimal outcomes. The experiments involving the PCA and LDA algorithms shared similar controlled variables due to the resemblance between the two methods, as both algorithms execute classification operations through dimensionality reduction techniques. Furthermore, in the classification experiments with the SVM algorithm, utilizing SVM in isolation led to unsatisfactory results. As a resolution, a fusion of SVM with PCA and LDA algorithms was implemented, resulting in the attainment of the highest performance. All findings of the traditional classification methods are displayed in Table 2.

Table 2. Summary of results.

Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
BP	71.99	87.74	79.09	80.89
k-NN	75.78	92.43	83.28	84.46
PCA	58.70	94.89	72.53	77.72
LDA	73.44	90.35	81.02	83.07
SVM	68.74	93.40	79.19	81.58

The classification efficiency results of traditional methods are shown in Table 3. The table data indicates that the k-NN and SVM algorithms are highly efficient, while the others have moderate performance.

Model	Training Duration (s)	Test Duration (s)	Total Duration (s)
BP	2378	12	2390
k-NN	178	11	189
PCA	338	11	349
LDA	569	12	581
SVM	127	11	138

Table 3. Classification efficiency results of traditional methods.

3.8. Comparative Summary of Results for Different Methods

Upon conducting experiments, we compared convolutional neural network (CNN) models with traditional methods, followed by an analysis of the results based on precision, recall, f1-score, accuracy, and total time. The findings, presented in Table 4, were complemented by the corresponding remote sensing image shown in Appendix A, Figure A7. From the data in the table, it was observed that AlexNet and EfficientNet-V2 exhibited relatively high precision, while GoogleNet and ResNet demonstrated higher recall values. Additionally, ResNet and EfficientNet-V2 showed relatively high f1-score and accuracy.

Notably, when not considering precision, k-NN, and SVM displayed higher efficiency, with ResNet standing out for higher efficiency when precision was taken into account.

Model	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	Total Time
AlexNet	96.81	87.57	91.96	91.58	846
VGG	90.98	87.98	89.46	89.41	3427
GoogleNet	87.10	95.23	90.98	90.99	1478
ResNet	92.47	94.29	93.37	93.27	698
EfficientnetV2	96.35	91.44	93.83	93.47	4611
BP	71.99	87.74	79.09	80.89	2390
k-NN	75.78	92.43	83.28	84.46	189
PCA	58.70	94.89	72.53	77.72	349
LDA	73.44	90.35	81.02	83.07	581
SVM	68.74	93.40	79.19	81.58	138

 Table 4. Summary table comparing results of different methods.

4. Discussion

In this study, we employed satellite remote sensing images obtained and preprocessed through the Google Earth Engine (GEE) platform to achieve walnut image classification. Both CNN models and traditional image classification methods were utilized to conduct experiments. By combining convolutional neural networks (CNNs) with remote sensing data, we aimed to enhance the accuracy and efficiency of walnut image classification [28].

In recent years, numerous researchers have conducted studies using remote-sensing images to examine various aspects of walnuts. For example, Albatrni et al. offered a comprehensive review of walnut shell adsorbents [52]. Ji et al. studied the genetic inheritance of walnuts and related factors [53]. Madrid et al. conducted real-time detection research on walnuts [54]. Hao Fei et al. studied cotton classification by employing multiple features and the random forest feature selection algorithm [55]. The ongoing expansion and shift towards large-scale farming in walnut planting areas have created a growing need for accurate and efficient methods for monitoring and analyzing cultivation areas. Despite the increasing scale of agricultural development, research on the use of satellite remote sensing methods for monitoring walnut planting areas remains relatively scarce. Traditional manual statistical methods, previously used for this purpose, have become impractical in light of the expanding cultivation areas. However, the utilization of satellite remote sensing techniques has emerged as an effective solution to address the challenges of monitoring and analyzing walnut planting areas in the context of the evolving agricultural landscape [56].

In this study, we selected BP, k-NN, PCA, LDA, and SVM as the traditional methods, considering their well-established nature and proven effectiveness in addressing a variety of problems. These algorithms have a demonstrated track record of delivering reliable results across diverse applications, as evidenced by previous research findings. For example, Yanbiao Xi et al. achieved a maximum accuracy of 84.19% by utilizing Sentinel-2 satellite data and various machine-learning algorithms for detailed tree species classification [57]. Rui He et al. also attained commendable accuracy, reaching 85%, by exploring the feasibility of tree species classification in the southwestern province of Sichuan, China, using Landsat satellite composite imagery [58]. To improve and train our models, we selected classic and versatile convolutional neural network (CNN) models such as AlexNet, VGG, GoogleNet, ResNet, and EfficientNetV2, covering a wide range of research in CNNs, including lightweight convolutional neural networks [59]. During the training process, we made specific adjustments to AlexNet, VGG, and EfficientNetV2 models. For AlexNet, we added convolutional and pooling layers and adjusted corresponding parameters. Similarly, we introduced dropout layers to prevent overfitting in the VGG model, leading to the relevant results. For EfficientNetV2, we mainly added Fuse-MBConv and MBConv modules, yielding relatively superior experimental outcomes. However, more adjustments were made to the GoogleNet and ResNet models. In GoogleNet, we modified the structure within

the inception module, added convolutional and pooling layers, and incorporated dropout layers to address overfitting, achieving satisfactory results. Similarly, in ResNet, we made modifications and additions to the sequential model structure, ultimately achieving positive outcomes. Xingrong Li et al. identified abandoned jujube orchards with an accuracy of 91.1% using multi-temporal high-resolution imagery and machine learning techniques, showing a significant enhancement in accuracy compared to other moderate-resolution satellite images [60]. However, our experimental approach, employing CNN-based models, clearly demonstrated superior performance compared to these established traditional methods and previous research findings.

The inaccurate classification of walnut areas in this study can be attributed to several reasons. Firstly, the low resolution of Landsat-8 imagery posed a significant challenge in a large-scale agricultural setting. While the use of the Sentinel-2 satellite did not yield satisfactory results due to issues with the satellite's resolution, it is important to note that this outcome does not necessarily imply that higher-resolution imagery would have no impact on improving accuracy. Higher-resolution images contain more detailed information and have a higher revisit frequency, providing more choices for research data. Therefore, from a comprehensive perspective, higher accuracy can be achieved through appropriate noise reduction processes. Furthermore, the study relied on robust and generalizable algorithms. Exploring and fine-tuning algorithms that are better suited to address the specific problem from an algorithmic perspective is expected to enhance accuracy. Additionally, leveraging diverse data sources such as unmanned aerial vehicle data, ground-based hyperspectral data, and image data could be considered from the standpoint of precision agriculture and smart farming. However, it is important to note that in addressing the issue of improved accuracy, the trade-off between data volume and model complexity must be carefully considered to mitigate the potential impact on the experiment's efficiency.

In this study, the classification efficiency of traditional methods was found to be higher than that of various convolutional neural network (CNN) models based on the experimental results. However, due to the large number of farm images to be detected for the classification problem addressed, the accuracy of walnut orchard detection becomes more crucial. Notably, throughout the research process, it was consistently observed that various CNN models achieved significantly higher accuracy than traditional classification methods. Moreover, they demonstrated superior performance in precision and f1-score parameters. Although in some cases, the recall values were relatively higher for the CNN models compared to traditional methods, when these values were accompanied by relatively low precision and f1-score values, it indicated relatively poorer performance. Thus, it is evident that in the classification of walnut remote sensing images, convolutional neural networks outperformed traditional methods.

Upon analyzing various convolutional neural network models, it becomes apparent that ResNet and EfficientNet-V2 demonstrate comparatively higher accuracy. Conversely, AlexNet and EfficientNet-V2 exhibit relatively higher precision, while GoogleNet and ResNet show relatively higher recall. In addition, ResNet and EfficientNet-V2 also display relatively higher f1-scores. Consequently, it can be deduced that ResNet and EfficientNet-V2 outperform other models in the classification of walnut remote sensing images. When considering efficiency, although EfficientNet-V2 achieves faster training speeds, it necessitates more training iterations to attain optimal accuracy, thereby resulting in a longer total training time in this study. Notably, this study integrated a substantial number of Fuse-MBConv and MBConv modules into the EfficientNet model to ensure detection accuracy, leading to a decreased efficiency of the lightweight algorithms. However, reducing these modules significantly resulted in a notable drop in accuracy. Thus, in this study, a significant reduction in the number of modules in EfficientNet-V2 while maintaining accuracy was unattainable. Consequently, ResNet outperformed EfficientNet-V2 in terms of effectiveness in this study.

A comprehensive comparative analysis between CNNs and conventional algorithms demonstrates the superior performance of CNNs, affirming the prevailing consensus within the field. Despite this, the efficiency aspect should be acknowledged, as traditional algorithms are generally favored, particularly in scenarios where precision requirements are not paramount. However, recent research endeavors have focused on developing lightweight CNNs aimed at expediting processing speed, reflecting the growing interest in addressing efficiency concerns associated with CNNs [61,62]. It is essential to note that this study utilized Landsat-8 satellite imagery at a 30-m resolution, which is widely suitable and well-suited for remote sensing applications in extensive agricultural landscapes. Nevertheless, challenges may arise when identifying small-scale agricultural plots with this resolution. Moreover, the utilization of higher-resolution satellite imagery has the potential to increase the computational complexity of corresponding algorithms, with uncertain impacts on detection efficiency. Furthermore, this study primarily concentrated on the integration and examination of core algorithmic components of lightweight models, such as EfficientNet-V2, without delving into detailed adjustments within these modules. As part of our future research agenda, we intend to explore refinements within these model modules to enhance detection efficiency while maintaining precision standards.

This investigation has demonstrated that the implementation of the ResNet model has resulted in comparable efficiency to traditional algorithms while maintaining a high level of precision. Furthermore, our approach has shown notably superior performance in contrast to classification outcomes obtained from other deep learning algorithms, thus enabling effective and accurate large-scale crop classification. The significance of this achievement is widespread, carrying implications that are relevant for both practical applications and prospective research endeavors. It is important to highlight that the use of satellite remote sensing imagery in this study has enabled the rapid acquisition of comprehensive and timely information, surpassing the capabilities of ground-based remote sensing. This attribute makes it highly suitable for agricultural detection and assessment purposes. Furthermore, the application of such data extends its utility beyond governmental and corporate realms to encompass a promising role in advancing individual pursuits, particularly in tasks related to detection, regulation, and related domains.

Remote sensing imagery has been widely used in previous research on walnut analysis. In these studies, researchers focused on extracting walnut planting areas and conducting growth analysis and monitoring using vegetation indices in conjunction with Google Earth Engine (GEE). These studies primarily emphasized the investigation of time series, feature bands, and vegetation indices [35]. In this study, we have taken the approach of integrating remote sensing imagery, leveraging convolutional neural networks (CNNs) and machine learning techniques to extract image features and achieve effective classification outcomes. There is potential for future research to combine these methods to achieve more efficient and precise detection. Furthermore, our future research aims to explore the acquisition of higher-resolution satellite imagery or resampling existing imagery, prioritizing efficiency while maintaining accuracy.

By observing and analyzing the experimental results, it is evident that the experiment successfully demonstrated the effectiveness of convolutional neural networks in walnut classification. This method not only reduces the burden on agricultural workers and improves classification efficiency but also provides a scientific basis for precision agriculture management. Optimizing agricultural resource utilization is expected to increase walnut yield and quality, thus generating positive impacts on the economy and ecology. Furthermore, in future research, there should be a focus on improving model performance and considering the integration of multi-source remote sensing data to enhance classification accuracy and reliability.

5. Conclusions

Amidst the era of expanding large-scale and mechanized agriculture, traditional manual census methods are encountering growing challenges attributable to time and labor constraints. The utilization of satellite remote sensing imagery in this study has emerged as a vital tool, facilitating the rapid acquisition of comprehensive and timely information, surpassing the capabilities of ground-based remote sensing. This characteristic makes it highly feasible for agricultural detection and assessment purposes, demonstrating its potential to address the limitations of traditional methods. Furthermore, the application of such data extends its utility beyond governmental and corporate spheres to encompass a promising role in advancing individual pursuits, particularly in tasks related to detection, regulation, and related domains. The successful combination of CNNs with satellite remote sensing for precise walnut classification in this study offers valuable insights into addressing these challenges and provides potential solutions to enhance agricultural monitoring and management.

This study is a significant interdisciplinary endeavor in the realms of agriculture and forestry as it has harnessed the capabilities of remote sensing and deep learning technologies to achieve precise classification of walnuts. This accomplishment notably enhances the automation level in agricultural and forestry production, thereby reducing labor costs and improving production efficiency. The precise classification of walnut trees harmonizes seamlessly with agricultural cycles, coupled with the inherent periodicity of satellite imagery. As a result, farmers gain the ability to enhance land and resource management, thereby reducing waste and promoting agricultural sustainability. The research leveraged high-resolution multispectral remote sensing images provided by Landsat-8 satellite, yielding valuable information pertaining to land cover and vegetation health, which contributes to achieving accurate classification and further propels the development of precision agriculture and smart farming. Furthermore, the study reinforced the utilization of deep learning models, including CNN, and has achieved commendable results, proving highly effective in the classification of remote sensing images in the agricultural domain. Through practical application, this approach adeptly captures fine-grained vegetation features, thereby enhancing classification accuracy. Moreover, the applicability of this method extends beyond walnut classification, showcasing its versatility in the classification and monitoring of various crops and offering guidance and inspiration for related research endeavors in the field.

Author Contributions: Conceptualization, J.W., Z.S. and X.L.; methodology, J.W. and X.L.; validation, J.W., Z.S. and S.L.; formal analysis, X.L., J.W. and K.H.; investigation, J.W., Z.S., S.L. and K.H.; resources, T.B. and Z.S.; data curation, J.W., X.L. and K.H.; writing—original draft preparation, J.W., Z.S. and X.L.; writing—review and editing, X.L. and T.B.; supervision, T.B., X.L. and Z.S.; funding acquisition, T.B. and X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Oasis Ecological Agriculture Corps Key Laboratory Open Project (202002), the Corps Science and Technology Program (2021CB041, 2021BB023, and 2021DB001), the Tarim University Innovation Team Project (TDZKCX202306 and TDZKCX202102), and the National Natural Science Foundation of China (61563046).

Data Availability Statement: The data used in the experiments are available upon request by contacting the authors.

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

Appendix A



Figure A1. Examples of partially labeled images.



Confusion Matrix for Test Data

Figure A2. AlexNet model confusion matrix.



Figure A3. VGG model confusion matrix.



Predicted Labels





Figure A5. ResNet model confusion matrix.



Figure A6. EfficientnetV2 model confusion matrix.





References

- 1. Banks, T. Property rights reform in rangeland China: Dilemmas on the road to the household ranch. *World Dev.* **2003**, *31*, 2129–2142. [CrossRef]
- Fu, X.; Niu, H. Key technologies and applications of agricultural energy internet for agricultural planting and fisheries industry. *Inf. Process. Agric.* 2022, 10, 416–437. [CrossRef]
- 3. Valjarević, A.; Algarni, S.; Morar, C.; Grama, V.; Stupariu, M.; Tiba, A.; Lukić, T. The coastal fog and ecological balance for plants in the Jizan region, Saudi Arabia. *Saudi J. Biol. Sci.* **2023**, *30*, 103494. [CrossRef] [PubMed]
- 4. Ramos, D.E. Walnut Production Manual; UCANR Publications: San Diego, CA, USA, 1997; Volume 3373.

- 5. Martínez, M.L.; Labuckas, D.O.; Lamarque, A.L.; Maestri, D.M. Walnut (*Juglans regia* L.): Genetic resources, chemistry, by-products. *J. Sci. Food Agric.* 2010, 90, 1959–1967. [CrossRef] [PubMed]
- 6. Bernard, A.; Lheureux, F.; Dirlewanger, E. Walnut: Past and future of genetic improvement. *Tree Genet. Genomes* **2018**, *14*, 1. [CrossRef]
- 7. Tian, J.; Wu, Y.; Wang, Y.; Han, F. Development and prospects of the walnut industry in China. In *VI International Walnut Symposium 861*; ISHS: Leuven, Belgium, 2009; pp. 31–38.
- 8. Fang, G.; Chen, Y.; Li, Z. Variation in agricultural water demand and its attributions in the arid Tarim River Basin. *J. Agric. Sci.* **2018**, *156*, 301–311. [CrossRef]
- 9. Hu, M.; Tian, C.; Zhao, Z.; Wang, L. Salinization causes and research progress of technologies improving saline-alkali soil in Xinjiang. J. Northwest A F Univ.-Nat. Sci. Ed. 2012, 40, 111–117.
- 10. Zhou, B.; Yang, L.; Chen, X.; Ye, S.; Peng, Y.; Liang, C. Effect of magnetic water irrigation on the improvement of salinized soil and cotton growth in Xinjiang. *Agric. Water Manag.* **2021**, *248*, 106784. [CrossRef]
- Liu, M.; Yang, J.; Li, X.; Mei, Y.; Jin, W. Effects of irrigation water quality and drip tape arrangement on soil salinity, soil moisture distribution, and cotton yield (*Gossypium hirsutum* L.) under mulched drip irrigation in Xinjiang, China. J. Integr. Agric. 2012, 11, 502–511. [CrossRef]
- Hou, X.; Xiang, Y.; Fan, J.; Zhang, F.; Hu, W.; Yan, F.; Xiao, C.; Li, Y.; Cheng, H.; Li, Z. Spatial distribution and variability of soil salinity in film-mulched cotton fields under various drip irrigation regimes in southern Xinjiang of China. *Soil Tillage Res.* 2022, 223, 105470. [CrossRef]
- 13. Feng, X.; Zhou, H.; Zulfiqar, S.; Luo, X.; Hu, Y.; Feng, L.; Malvolti, M.E.; Woeste, K.; Zhao, P. The phytogeographic history of common walnut in China. *Front. Plant Sci.* 2018, *9*, 1399. [CrossRef]
- 14. Baojun, Z.; Yonghong, G.; Liqun, H. Overview of walnut culture in China. In *VI International Walnut Symposium* 861; ISHS: Leuven, Belgium, 2009; pp. 39–44.
- 15. Shigaeva, J.; Darr, D. On the socio-economic importance of natural and planted walnut (Juglans regia L.) forests in the Silk Road countries: A systematic review. *For. Policy Econ.* **2020**, *118*, 102233. [CrossRef]
- 16. Dong, Y.-Z.; Liang, F.-L.; Wang, Z.-Y.; Xie, E.-J.; Zhu, X.-H. Investigation and analysis on the wild walnut in Gongliu, Xinjiang. *J. Plant Genet. Resour.* **2012**, *13*, 386–392.
- 17. Gu, X.; Zhang, L.; Li, L.; Ma, N.; Tu, K.; Song, L.; Pan, L. Multisource fingerprinting for region identification of walnuts in Xinjiang combined with chemometrics. *J. Food Process Eng.* **2018**, *41*, e12687. [CrossRef]
- 18. Lv, X.; Ming, D.; Chen, Y.; Wang, M. Very high resolution remote sensing image classification with SEEDS-CNN and scale effect analysis for superpixel CNN classification. *Int. J. Remote Sens.* **2019**, *40*, 506–531. [CrossRef]
- 19. Ghassemian, H. A review of remote sensing image fusion methods. Inf. Fusion 2016, 32, 75-89. [CrossRef]
- Cheng, G.; Han, J.; Lu, X. Remote sensing image scene classification: Benchmark and state of the art. *Proc. IEEE* 2017, 105, 1865–1883. [CrossRef]
- Bo, Y.; Wang, H. The application of cloud computing and the internet of things in agriculture and forestry. In Proceedings of the 2011 International Joint Conference on Service Sciences, Taipei, Taiwan, 25–27 May 2011; pp. 168–172.
- Zhang, S.; Zhang, C.; Ding, J. Disease and insect pest forecasting model of greenhouse winter jujube based on modified deep belief network. *Trans. Chin. Soc. Agric. Eng.* 2017, 33, 202–208.
- 23. Li, W.; Fu, H.; Yu, L.; Cracknell, A. Deep learning based oil palm tree detection and counting for high-resolution remote sensing images. *Remote Sens.* 2016, 9, 22. [CrossRef]
- 24. Campbell, J.B.; Wynne, R.H. Introduction to Remote Sensing; Guilford Press: New York, NY, USA, 2011.
- 25. McCarthy, N.F.; Tohidi, A.; Aziz, Y.; Dennie, M.; Valero, M.M.; Hu, N. A Deep Learning Approach to Downscale Geostationary Satellite Imagery for Decision Support in High Impact Wildfires. *Forests* **2021**, *12*, 294. [CrossRef]
- Iban, M.C.; Sahin, E. Monitoring land use and land cover change near a nuclear power plant construction site: Akkuyu case, Turkey. *Environ. Monit. Assess.* 2022, 194, 724. [CrossRef] [PubMed]
- Zhang, C.; Zhang, H.; Tian, S. Phenology-assisted supervised paddy rice mapping with the Landsat imagery on Google Earth Engine: Experiments in Heilongjiang Province of China from 1990 to 2020. *Comput. Electron. Agric.* 2023, 212, 108105. [CrossRef]
- Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 2017, 202, 18–27. [CrossRef]
- Dillon, T.; Wu, C.; Chang, E. Cloud computing: Issues and challenges. In Proceedings of the 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, Australia, 20–23 April 2010; pp. 27–33.
- Kumar, L.; Mutanga, O. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sens.* 2018, 10, 1509. [CrossRef]
- 31. Tamiminia, H.; Salehi, B.; Mahdianpari, M.; Quackenbush, L.; Adeli, S.; Brisco, B. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* **2020**, *164*, 152–170. [CrossRef]
- 32. Voulodimos, A.; Doulamis, N.; Doulamis, A.; Protopapadakis, E. Deep learning for computer vision: A brief review. *Comput. Intell. Neurosci.* 2018, 2018, 7068349. [CrossRef]

- 33. Jung, J.; Maeda, M.; Chang, A.; Bhandari, M.; Ashapure, A.; Landivar-Bowles, J. The potential of remote sensing and artificial intelligence as tools to improve the resilience of agriculture production systems. *Curr. Opin. Biotechnol.* **2021**, *70*, 15–22. [CrossRef]
- Cheng, G.; Xie, X.; Han, J.; Guo, L.; Xia, G.-S. Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2020, 13, 3735–3756. [CrossRef]
- 35. Shi, Z.; Zhang, R.; Bai, T.; Li, X. Walnut Acreage Extraction and Growth Monitoring Based on the NDVI Time Series and Google Earth Engine. *Appl. Sci.* **2023**, *13*, 5666. [CrossRef]
- 36. Wang, Y.; Li, G.; Wang, S.; Zhang, Y.; Li, D.; Zhou, H.; Yu, W.; Xu, S. A Comprehensive Evaluation of Benefit of High-Standard Farmland Development in China. *Sustainability* **2022**, *14*, 10361. [CrossRef]
- 37. Valjarević, A.; Djekić, T.; Stevanović, V.; Ivanović, R.; Jandziković, B. GIS numerical and remote sensing analyses of forest changes in the Toplica region for the period of 1953–2013. *Appl. Geogr.* **2018**, *92*, 131–139. [CrossRef]
- Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
- 39. O'Shea, K.; Nash, R. An introduction to convolutional neural networks. *arXiv* 2015, arXiv:1511.08458.
- Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* 2012, 25, 1–9. [CrossRef]
- Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Van Esesn, B.C.; Awwal, A.A.S.; Asari, V.K. The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv* 2018, arXiv:1803.01164.
- 42. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
- 43. Sengupta, A.; Ye, Y.; Wang, R.; Liu, C.; Roy, K. Going deeper in spiking neural networks: VGG and residual architectures. *Front. Neurosci.* **2019**, *13*, 95. [CrossRef]
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
- 45. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2016; pp. 770–778.
- Wu, Z.; Shen, C.; Van Den Hengel, A. Wider or deeper: Revisiting the resnet model for visual recognition. *Pattern Recognit*. 2019, 90, 119–133. [CrossRef]
- 47. Tan, M.; Le, Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the International Conference on Machine Learning, PMLR, Virtual. 18–24 July 2021; pp. 10096–10106.
- 48. Martinez, A.M.; Kak, A.C. Pca versus Ida. IEEE Trans. Pattern Anal. Mach. Intell. 2001, 23, 228–233. [CrossRef]
- 49. Li, J.; Cheng, J.; Shi, J.; Huang, F. Brief introduction of back propagation (BP) neural network algorithm and its improvement. In *Advances in Computer Science and Information Engineering*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 2, pp. 553–558.
- 50. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]
- 51. Goutte, C.; Gaussier, E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In *European Conference on Information Retrieval*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 345–359.
- 52. Albatrni, H.; Qiblawey, H.; Al-Marri, M.J. Walnut shell based adsorbents: A review study on preparation, mechanism, and application. J. Water Process Eng. 2022, 45, 102527. [CrossRef]
- 53. Ji, F.; Ma, Q.; Zhang, W.; Liu, J.; Feng, Y.; Zhao, P.; Song, X.; Chen, J.; Zhang, J.; Wei, X. A genome variation map provides insights into the genetics of walnut adaptation and agronomic traits. *Genome Biol.* **2021**, *22*, 300. [CrossRef] [PubMed]
- 54. Madrid, R.; García-García, A.; Cabrera, P.; González, I.; Martín, R.; García, T. Survey of commercial food products for detection of walnut (Juglans regia) by two elisa methods and real time pcr. *Foods* **2021**, *10*, 440. [CrossRef] [PubMed]
- 55. Fei, H.; Fan, Z.; Wang, C.; Zhang, N.; Wang, T.; Chen, R.; Bai, T. Cotton Classification Method at the County Scale Based on Multi-Features and Random Forest Feature Selection Algorithm and Classifier. *Remote Sens.* **2022**, *14*, 829. [CrossRef]
- 56. Li, J.; Pei, Y.; Zhao, S.; Xiao, R.; Sang, X.; Zhang, C. A review of remote sensing for environmental monitoring in China. *Remote Sens.* **2020**, *12*, 1130. [CrossRef]
- 57. Xi, Y.; Ren, C.; Tian, Q.; Ren, Y.; Dong, X.; Zhang, Z. Exploitation of time series sentinel-2 data and different machine learning algorithms for detailed tree species classification. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2021, 14, 7589–7603. [CrossRef]
- He, R.; He, B.; Shi, Y.; Chen, L.; Wang, Z.; Lai, X. Feasibility of Using Landsat Synthetic Images to Classify Tree Species in Southwest Sichuan, China. In Proceedings of the IGARSS 2022—2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 5692–5695.
- Zhou, Y.; Chen, S.; Wang, Y.; Huan, W. Review of research on lightweight convolutional neural networks. In Proceedings of the 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC), Chongqing, China, 12–14 June 2020; pp. 1713–1720.
- 60. Li, X.; Yang, C.; Zhang, H.; Wang, P.; Tang, J.; Tian, Y.; Zhang, Q. Identification of Abandoned Jujube Fields Using Multi-Temporal High-Resolution Imagery and Machine Learning. *Remote Sens.* **2021**, *13*, 801. [CrossRef]

- Anisimov, D.; Khanova, T. Towards lightweight convolutional neural networks for object detection. In Proceedings of the 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Lecce, Italy, 29 August–1 September 2017; pp. 1–8.
- 62. Wieczorek, M.; Siłka, J.; Woźniak, M.; Garg, S.; Hassan, M.M. Lightweight convolutional neural network model for human face detection in risk situations. *IEEE Trans. Ind. Inform.* 2021, *18*, 4820–4829. [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.