Analysis of Image Data Augmentation in Smart Agriculture

Mudong Liu *

Department of Mathematics, The University of Manchester, Manchester M139PL, United Kingdom

* Corresponding Author Email: mudong.liu@postgrad.manchester.ac.uk

Abstract. The demand for agricultural production automation in smart agriculture has made deep learning widely used in tasks, e.g., disease identification, weed-crop segment and fruit detection based on image processing. However, in agricultural production environments, there are limitations in data acquisition, so data enhancement technology has also been widely used. This study introduces two data enhancement methods commonly used in smart agriculture, namely image transformation and GAN-based, and analyzes their specific applications and roles in three common task scenarios: disease identification, weed-crop segment, and fruit detection. According to the analysis, the data enhancement is mainly used to deal with class imbalance problems and over-fitting problems, and can significantly improve model performance in specific scenarios; the complexity of tasks and image structures has an impact on the application effects of different methods, and image segmentation tasks. It is necessary to increase the number of image modes, so methods such as GAN-based are adopted. However, due to the complexity of the image, fruit detection uses simple transformation methods to achieve better results. In addition, data enhancement methods designed for specific tasks, such as simulating synthetic images, can also achieve good results. However, data augmentation technology still has mode collapse problems and is too cumbersome in its application, which points out the direction for subsequent research. Overall, this research sorts out the applications of common data enhancement methods, analyzes the characteristics and limitations of these methods, and hopes to provide guidance for subsequent research.

Keywords: Smart Agriculture, image transformation, GAN-based, data augmentation technology.

1. Introduction

Since the start of 21st century, as the contradiction between the growing world population and limited arable land develops, as well as the concept of sustainable development got popular, more challenges are raised to agricultural production in the worldwide: larger-scale production, refined management, and more environmentally friendly production method [1]. Specifically, this requires more refined control of crop growing conditions in agricultural production, real-time monitoring of possible pests, diseases, and weeds, the use of better-quality seeds, and more. At the same time, in larger-scale production, sowing, picking and other procedures also need to be automated. It's too much work for human workers to perform these operations manually, and it is also difficult to avoid errors that can cause potential losses. In technique aspects, although sensors allow physical and chemical indicators in the crop growth environment to be converted into electrical signals to achieve digital management, tasks such as crop pest and disease identification, agricultural weed identification, and automatic fruit identification and picking still require judgment based on visual features in the environment. In other words, these tasks involve image processing. Henec, in the early years, they could only be completed manually and could not be easily processed with computer technology.

In recent years, deep learning technology has achieved rich results in the field of computer vision and has also been applied in many other industries. Deep learning models headed by CNN have powerful image processing capabilities: such models can extract and learn specific hidden features from pixel-based images, and perform classification or regression prediction operations on the images based on these features. In the field of agriculture, in recent years, there has been a lot of research on using this type of image recognition model for tasks such as detection and classification of crop leaf diseases and insect pests, identification of crop growth status, identification and control of weeds, and identification of crop fruits.

In such computer vision tasks based on deep learning, a large amount of high-quality annotated data is often required to train the model. However, for specific tasks, especially in the agricultural field, most images have problems such as cluttered backgrounds and indistinct recognition subjects, making it difficult to obtain a large amount of high-quality image data [2]. In this case, model training can only be performed with a limited-size data set. However, since deep learning models often contain many parameters, insufficient sample size can easily lead the model to learning overly specific features, which results in overfitting. This results in reduced model generalization performance and poor performance in identifying targets in new images in practical applications. In addition, for image classification tasks, the data set may also suffer from class imbalance. For example, in the identification of leaf diseases, for a specific plantation environment, among many reliable crop leaf images, it is likely that only a few images contain leaves infected with a specific type of disease. During model training, this may cause the model to predict every input simply to the classes that the amount in training set is the largest, since in this strategy the accuracy has been high enough.

In this case, a simple and effective processing method is data enhancement. That is, through image transformation and other means, the data set is expanded from limited original data. Data enhancement can obtain more different data at the pixel level while ensuring that the basic visible features of the image remain unchanged. Through the expanded data set, the model can learn more general features and avoid overfitting of the model. This paper will review the application of data enhancement technology in smart agriculture and try to influence future research directions. This rest part of the paper mainly consists of 6 parts. Section 2 introduces common data augmentation methods, Sections 3 to 5 introduce the application of data augmentation technology in three typical problems in smart agriculture, and analyzes its motivations and reasons for using data augmentation. Section 6 analyzes the current shortcomings and possible future research directions of data augmentation methods. Section 7 summarizes the current research and application status of data augmentation in smart agriculture.

2. Methodology

This section introduces commonly used data augmentation methods. At present, common data enhancement methods mainly transform original images. With the development of deep learning technology, virtual image sample generation methods based on GAN and its derivative networks have gradually been developed and applied. The original image transformation method includes geometric transformation and texture transformation of the original image. By applying transformation to the original image, a new image can be obtained. These images often have similar key visual features to the original images, so the dataset can be expanded by adding labels consistent with the original images and then adding them back to the original dataset. Geometric transformation is the most basic data enhancement method, including stretching, skew, scaling, rotation, translation, flipping and other geometric transformations, as well as methods such as direct cropping or erasing of images. At the pixel level, these geometric transformations only change the position of the pixels in the image, but do not directly affect the pixel values of each part of the image. Texture transformation includes noise augmentation and blur augmentation, and a new image is obtained by performing operations on the pixel values of the entire image. Noise augmentation uses a noise matrix where each element is randomly generated from a Gaussian distribution. Superimpose the values of this matrix directly onto the pixel matrix of the original image to obtain the enhanced image. The Gaussian distribution of the noise matrix is generated, and its mean and variance usually have small values, so the visual effect of the transformed image and the original image is not much different. However, after superimposing Gaussian noise, the new image is a completely new model for model training, so it can also enhance the training effect. Fuzzy augmentation updates the value of each pixel based on the values of surrounding pixels. The update algorithm can be a smoothing algorithm, such as taking the median or average. This transformation will visually blur and distort the image, but will not change other visual characteristics such as the overall tone.

In addition to these two types of transformations, there are also methods for optical transformation of images, including changing the brightness, contrast, and sharpness of the image, and directly transforming the color space. However, these methods will have a significant impact on some visual characteristics of the image, especially color space transformation. This type of transformation cannot be used if color features are important visual features of the data being analyzed. For such reasons, these methods are rarely used in actual research. In addition, different forms of transformations can be freely combined to process images. In general, the image transformation method is simple to operate and has been integrated into the mainstream deep learning frameworks Tensorflow and PyTorch. Related methods can be directly called for data enhancement.

On the other hand, virtual image sample generation methods based on GAN and its derivative network uses GAN or its derivative model to learn the implicit distribution in the original data set, and then uses its generator to use random noise as input to learn the original data from the original data. Generate new augmented images from the set distribution. GAN, the generative adversarial network, proposed and developed by Goodfellow [3], consists of two parts: a generator G and a discriminator D. The goal of the generator G is to generate fake data that is realistic enough. It receives randomly generated noise input and generates fake samples. The goal of the discriminator D is to determine as accurately as possible whether the input data is generated by the generator or real data. It also receives real samples and fake samples generated by G, and output the judgment results of the authenticity of the picture. The training of GAN is completed through the mutual confrontation between the two parts to form a dynamic balance: in order to obtain a high-quality augmented image, D should be unable to distinguish between the image generated by G and the real image. At the same time, D should also have as strong an image as possible. The ability to distinguish authenticity. The loss function V of GAN and its optimization have the following forms:

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log D(x)] + E_{z\sim p_{data}(z)}[\log(1 - D(G(z)))] \quad (1)$$

Here, $D(\cdot)$ and $G(\cdot)$ are the outputs of the discriminator and generator respectively, while $x\sim p_{data}(x)$ and $z\sim p_{data}(z)$ respectively represent the real data and random noise and their distribution in the model input. In actual training, the parameters of D are first fixed, and the maximum value of V is found, i.e., the loss of D is maximized, thereby optimizing the parameters of the generator G; then, the image generated by the optimized G and the real image are used as input, minimize V and update the parameters of D, thus completing an iterative training cycle.

3. Disease Identification

This section introduces the application of data enhancement in the identification of crop diseases and pests. Both image transformation and GAN-based methods can be applied in this task. Scholars applied MobileNet and self-build CNN to potato leaf disease identification [4]. According to different disease conditions and visual characteristics, potato leaves are divided into three categories: early bright, late bright and healthy. However, the leaf images in the healthy category are only 15% of the other two categories, that is, the data set has a category imbalance. To this end, the method of rotating, translating, scaling, and flipping the healthy leaf images was used for data enhancement, and the healthy class was expanded to the same size as the other two classes, thus obtaining an expanded data set. The original data set and the expanded data set were used to train self-build CNN respectively, and the expanded data set was used to train MobileNet that applied transfer learning technology as a control. The results show that compared to not using data enhancement, the self-build CNN trained using the expanded data set has a closer training loss and validation loss, which shows that there is no overfitting phenomenon during the training process. At the same time, based on the confusion matrix, after applying data enhancement, the overall misclassification rate was reduced from 37% to 9%, reaching an accuracy close to MobileNet.

Others used deep convolutional networks to classify olive leaf diseases [5]. The images used in the study were taken directly by mobile phones, with a total of 3,400 pictures. At the same time, with the assistance of agricultural engineers, these images were divided into three categories: leaf images infected with Aculus olearius, images infected with Olive peacock spot and the healthy leaves. However, the data collected also suffer from class imbalance. This imbalance may have a negative impact on the training of the model, so similarly, data augmentation is performed on classes with smaller sample sizes in the dataset. The researchers used image transformation methods such as zooming, brightness, horizontal flip, and rotation to complete data enhancement. These methods are integrated into the ImageDataGenerator class of the Keras framework. The original data and the augmented data were used to train self-build CNN, pre-trained VGG16 and VGG19 respectively to form a comparative experiment. It turns out that without using data augmentation, the validation loss will get higher and higher as epochs increase, while this result does not occur when data augmentation is used. This shows that data augmentation prevents the model from overfitting.

GAN-based data enhancement methods can also be used for disease identification tasks. Also, for the tomato leaf disease classification task, previous study uses deep convolutional generative adversarial network (DCGAN) for data enhancement [6]. Compared with GAN that uses a simple fully connected network structure, technologies with stronger image processing capabilities such as convolution structure and batch normalization are introduced into DCGAN, and the ReLU function and LeakyReLU function are used as the activation function of the middle layer, which makes DCGAN Both the convergence speed and the quality of the generated images are greatly improved. Compared with traditional data enhancement methods, the images generated by DCGAN can be clearly divided into several categories in terms of distribution, and have visual features that are more similar to the original images. In conclusion, both image transformation and GAN-based methods are available in this disease identification task, and GAN-based methods tends to achieve better results. And they are widely used in related research.

4. Crop-Weed Segment

This section introduces the application of data enhancement in the crop-weed segment task. As plants that provide no economic value, weeds compete with crops for nutrients, water and sunlight as they grow, thus negatively impacting agricultural output. Therefore, regular inspection and removal of weeds is an important step in traditional agriculture. In modern smart agriculture, the core of the problem is to identify weeds in time. This is an image segmentation task that requires segmenting weeds in an image that is a mixture of weeds and crops, and attaching labels to the pixels in the corresponding areas. However, due to the small area of weed leaves and the fact that different plants in the image may block each other, labeling the image requires a lot of time and labor costs. Therefore, when collecting data for research by yourself, usually only a small amount of annotated data can be obtained. To meet the needs of model training, one designed a data enhancement method based on foreground and background to synthesize the images in the U-Net pre-training stage: using the weeds in the annotated data as the foreground, and splicing it with the background image without weeds, you can synthesize images containing semantic annotations [7]. In this way, the data set was expanded from 800 images to 10,000 images. The results show that this effectively alleviates the overfitting phenomenon during model training.

In addition, the weed-crop segment task also requires the ability to identify crops and weeds of different types and growth stages under different weather and other background environments. But in fact, due to the diversity of climate, crops, and weeds, it is difficult to collect enough data that can cover all modes, and the problem of class imbalance is unavoidable. Others proposed a data augmentation strategy for generating semi-artificial image data using DCGAN and cGAN [8]. The image generated by this strategy includes artificially generated crop objects and real background. The author compared this strategy with data augmentation strategies based on image and texture transformation: In the weed-crop image segmentation task based on the data in the Bonn sugar beet dataset, the original dataset with the same sample size was used, and only A dataset for image segmentation is trained using a dataset of images generated using image transformations or texture transformations and a dataset mixed with images synthesized using GANs. The process of generation is shown in Fig. 1 [8].

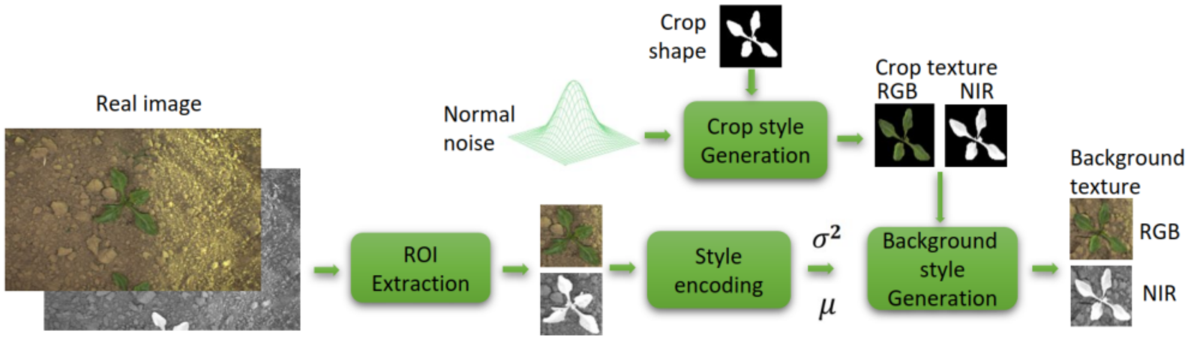

Figure 1. A process of GANs [8].

Taking Mean Intersection over Union (denoted as mIoU) as the evaluation metric, the results show that the model trained using data mixed with synthetic crop images generated by the proposed approach performed best. From the researches above, it's shown that in segment tasks, data augmentation is also motivated by the limited access to annotated data. However, both two research choose to generate images that has different mode rather than traditional image transformation. It also helps with improving the performance of model and dealing with overfitting problems.

5. Fruit Detection

This section introduces the application of data augmentation in fruit detection. Currently, there is a strong demand on the automation of the fruit processing in agriculture. Hence researches on identifying the fruits as well as estimating the quantity and quality of fruits are motivated recently. It also helps farmers to make better decision on stage of harvesting, which can reduce the cost of operation. And similarly, the using of data augmentation is motivated by the high cost of collecting and annotating samples.

To achieve this goal, [9] used the method of artificially synthesized images for data enhancement when studying tomato detection. Since the actual images used to detect fruits usually have soil and other plants as the background, a synthetic image consisting of randomly distributed brown, green, and red color patches is used to simulate the actual image. The brown and green color blocks simulate the soil and plant leaves in the background, the red color block simulates the tomato fruit, and a layer of Gaussian filter is added. The detection model of the Inception-ResNet architecture was trained using 24,000 synthetic images, and achieved an accuracy of 91% when tested using the real tomato picture sample.

Besides, data augmentation can be even used to improve the performance of model directly. [10] used flip, scale and PCA enhancement techniques when using R-CNN to detect apples, mangoes and almonds. PCA enhancement is a texture transformation by applying random perturbations to RGB intensity. Data augmentation is used in each training epoch to reduce the computational cost of random augmentation in large amounts of data. The results show that PCA has no significant impact on the training results, while flip and scale significantly improve model performance while also reducing the sample size required to achieve specific model performance. Among them, using flip and scale enhancement at the same time has the greatest improvement in model performance. In fruit detection task, it seems that the complexity of image prevents GAN-based data argumentation methods from application. However, simple idea such as synthetic simulating images or directly simple flip and scale transformation effectively improve the performance of models. In addition, sometimes mixture of original image and augmented image even performs better than dataset all consists of original data, even if there is no class imbalance.

6. limitations and prospects

At present, there are still some problems in the application and research of data enhancement methods in smart agriculture. The virtual image synthesis method still suffers from the mode collapse problem, that is, after the GAN is trained, the generator can only generate a single mode or a few modes. At this time, GAN can only generate many poor-quality images, almost losing the role of data enhancement. This may be because the data distribution is often more complex than the distribution that the complexity of GAN supports. At this time, GAN may converge to a smaller part of the actual distribution, leading to mode collapse. At the same time, the GAN type model itself also has high complexity, which increases the cost of data enhancement processing and is not conducive to practical applications. In addition, although image transformation and texture transformation methods are simple and easy to use, they sometimes have limited effects and cannot effectively avoid category imbalance and overfitting problems.

In this regard, future research can consider improving the network structure and exploring more effective methods to prevent mode collapse. Or introduce existing network structures such as WGAN that can improve mode collapse. In addition, you can also explore data enhancement models that consume less resources, or encapsulate simple GAN networks into existing deep learning frameworks to reduce the time and space costs required for data enhancement, achieve lightweight data enhancement technology, and increase Its application value in smart agriculture.

7. Conclusion

To sum up, this study summarizes common data augmentation methods in smart agriculture, and the application of data augmentation technology in three typical problems in smart agriculture research, disease identification, weed-crop segment, and fruit detection. This study compares the use and results of different data enhancement methods in different task scenarios, and attempts to explain the differences, in order to provide guidance for subsequent theoretical and empirical research. For these applications, this study summarizes the main characteristics of data enhancement methods in smart agriculture applications, and proposes the shortcomings and improvement directions of existing research. Data augmentation is usually used to solve class imbalance problems and over-fitting problems, and can improve model performance and convergence speed. In some tasks, this improvement is noticeable. Different tasks require different data enhancement methods. Too high image complexity will limit the use of GAN-based methods, and the complexity of the task will make it difficult for image transformation methods to play a significant role. The current GAN-based method still lacks the means to deal with mode collapse, and its complex structure will also affect its application value.

References

[1] M. Altalak, M. Ammad uddin, A. Alajmi and A. Rizg, Applied Sciences 12 (12), 5919 (2022).

[2] C. Shorten and T. M. Khoshgoftaar, Journal of Big Data, 6 (1) (2019).

[3] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al. Communications of the ACM 63 (11), 139 - 144 (2020).

[4] U. Barman, D. Sahu, G. G. Barman and J. Das, “Comparative assessment of deep learning to detect the leaf diseases of potato based on data augmentation,” In 2020 International Conference on Computational Performance Evaluation (IEEE, NY 2020) pp. 682 - 687.

[5] S. Uğuz and N. Uysal, Neural computing and applications 33 (9), 4133 - 4149 (2021).

[6] Q. Wu, Y. Chen and J. Meng, IEEE Access 8, 98716 - 98728 (2020).

[7] K. Zou, X. Chen, Y. Wang, C. Zhang, and F. Zhang, Computers and Electronics in Agriculture 187, 106242 (2021).
[8] M. Fawakherji, V. Suriani, D. D. Bloisi and D. Nardi, Weed Segmentation in Precision Farming 7 (5) (2022).
[9] M. Rahnemoonfar, and C. Sheppard, Sensors 17 (4), 905 (2017).
[10] S. Bargoti, and J. Underwood, “Deep fruit detection in orchards,” in 2017 IEEE International Conference on Robotics and Automation (IEEE, London, 2017) p 15.