

Training Data Augmentations for Improving Hyperbola Recognition in Ground Penetrating Radar B-Scan Image for Tree Roots Detection

Zeqing Li,^{a,b} Xiaowei Zhang,^{a,b} Haibin Li,^{a,b} Zepeng Wang,^{a,b} and Jian Wen^{a,b,*}

Improving the detection accuracy of hyperbola in B-scan images has been a considerable challenge for ground penetrating radar (GPR) to detect tree roots. In this paper, a method for data enhancement and target detection, both based on deep learning was proposed to identify hyperbolas in GPR B-scan images. First, the authors used a cyclic consistent adversarial network (CycleGAN) to augment the original data. In this procedure, the hyperbolic features of the images were preserved and created a wider variety of training samples. Then, the authors could apply the enhanced dataset to the YOLOv5 detection model to evaluate the effectiveness of their method. Meanwhile, the detection effects of Yolov3, Yolov5, Faster R-CNN, and CenterNet detection models on the enhanced dataset were compared. The results showed that applying the enhanced dataset to the Yolov5 detection model exhibited better detection accuracy compared to other combinations of datasets and detection models. The authors demonstrate that the proposed method increases data diversity and the number of samples, improving the precision and recall of hyperbolic curves. These results provide a new method for tree root localization with important effects.

DOI: [10.15376/biores.18.1.484-504](https://doi.org/10.15376/biores.18.1.484-504)

Keywords: Ground-penetrating radar (GPR); Data enhancement; YOLOv5; Cycle-Consistent Adversarial Networks (CycleGAN); Tree root detection

Contact information: a: School of Technology, Beijing Forestry University, Beijing 100083, P. R. China; b: Joint International Research Institute of Wood Nondestructive Testing and Evaluation, Beijing Forestry University, Beijing 100083, P. R. China; *Corresponding author: wenjian@bjfu.edu.cn

INTRODUCTION

As a vital organ for tree growth and development, roots have a significant effect on the entire life cycle of trees and also play an equally important role in the material cycle and energy flow of the soil (Gill and Jackson 2000; Reubens *et al.* 2007). The specific structure of the tree root system can be determined through intensive analysis of parameters such as root diameter, orientation, burial depth, root water content, and root distribution (Danjon and Reubens 2008). Detection of subsurface root distribution and study of the specifics of root parameters are often performed using excavation methods, monoliths, *etc.* (Guo *et al.* 2013; Riedell and Osborne 2017), which are complex to operate and can cause irreversible damage to the soil environment and trees. Therefore, the nondestructive detection of tree roots is a challenging task.

Ground penetrating radar (GPR), a nondestructive testing technique (NDT) has been widely used in tree roots research. Compared with traditional tree root NDT methods, such as ultrasonic pulse velocity (UPV) analysis (Wang and Li 2015; Sarro *et al.* 2021) and electrical resistance tomography (ERT) (LaBrecque and Yang 2001; Kemna *et al.*

2002), GPR has the advantages of high efficiency, safety, and high interference resistance (Alani *et al.* 2018; Mihai *et al.* 2019; Aboudourib *et al.* 2021). The method GPR can accurately detect targets based on the relative permittivity between the targets and other media. This is because the water content in the tree root system is vastly different from the water content in the adjacent soil (Pettinelli *et al.* 2014; Tanoli *et al.* 2019), resulting in a noticeable difference in relative permittivity. Therefore, it is feasible to use GPR to detect root structure. The use of ground penetrating radar to detect the size of plant roots is one of the important directions of current research. Zhou *et al.* (2019) proposed a model combining ground penetrating radar with electric field methods, which could automatically and quickly fit a hyperbola in the retained image area and could effectively obtain the depth and radius of buried objects. When the root system is detected, the general distribution of the root system can be observed visually, and further information such as the location, size and orientation of the root system can be obtained, and the specific structure of the root system of the tree can be established. This information can better analyze the growth and health of the tree, the roots usually appear as hyperbolas in the radar B-scan image. Automatic detection of hyperbolic features in the images can improve efficiency and recognition accuracy. Before the B-scan image hyperbolas can be identified, the image first has to be preprocessed. However, in the actual detection environment, because of the random nature of soil distribution, radar hardware, wave interactions, and different underground media (Daniel *et al.* 2016), there are various noises existing in the detected B-scan images. Therefore, B-scan images are often pre-processed by several signal and image processing methods. Wen *et al.* (2020) proposed a shearlet transform to perform noise removal from B-Scan images and achieved better denoising results in some image evaluation metrics. Deep learning models are being widely used to detect the internal structure of tree root systems (Xiang *et al.* 2019; Hou *et al.* 2021; Zhang *et al.* 2021). Hou *et al.* (2021) proposed using MS R-CNN architecture for the detection of GPR subsurface scanned objects, while using the transfer learning technique to obtain pre-trained models to solve the problem of the insufficient model training set (93 GPR root scans). Zhang *et al.* (2021) used the Faster R-CNN to train 1442 GPR B-scan images (282 for real images and 1160 for simulation images) to achieve automatic recognition and localization of hyperbolas in GPR images. It is not easy to perform effective automatic detection of tree root systems with these methods. The main obstacle is the single type and an insufficient number of datasets.

To overcome this obstacle, data augmentation was performed on the tree root system dataset. The traditional image augmentation of the tree root system uses scales, stretches, flips, crops, and obtains the GPR B-scan images, increasing the number of training data by changing the original data without expanding the number of a real dataset. The method of using GprMax 3.0 software (University of Edinburgh, Dr. Antonis Giannopoulos, UK) to generate simulation images to extend the dataset has been widely used (Todkar *et al.* 2021; Dewantara and Parnadi 2022), and the dataset composition has been expanded from real images to a joint composition of simulation images based on GprMax 3.0 and real images. This enabled the expansion of the quantity and variety of training data. However, the problem of the limited number of real datasets remains unresolved. Therefore, this paper used CycleGAN (Zhu *et al.* 2017) to transform the simulation images generated by GprMax3.0 into the corresponding generated images with high similarity to the real images, which achieved the augmentation of the real dataset. This method provides an effective expansion of real data with insufficient diversity and helps to improve the accuracy of target detection methods.

YOLOv3 and Faster R-CNN (Ren *et al.* 2017) with anchor-base are the most versatile target detection methods. YOLOv3 and Faster R-CNN have been successfully applied in agriculture (Liu *et al.* 2020; Thanh Le *et al.* 2021), geology (Ma *et al.* 2019; Davletshin *et al.* 2021), remote sensing (Zhou *et al.* 2019; Li *et al.* 2022), and medicine (Rosati *et al.* 2020; Yao *et al.* 2020). YOLOv3 is widely used in forestry fields, such as tree health classification (Yarak *et al.* 2021), forest census (Zheng *et al.* 2019), and tree species identification for detection. YOLO series continues to evolve and improve in both detection accuracy and detection speed. Recently, YOLOv5 and Anchor-free based CenterNet are being used in a wide variety of fields. In this study, the real data were augmented by CycleGAN to obtain a mixture dataset, and the three kinds of data in the mixture dataset were combined to obtain seven datasets, while comparing the effect of different datasets trained with YOLOv5. The results show that enhanced datasets have better training and recognition results. Meanwhile, the detection results of four models, YOLOv3, YOLOv5, Faster R-CNN, and CenterNet, were compared on the same dataset. The results show that YOLOv5 was highly accurate in detecting tree roots.

EXPERIMENTAL

Working Principle of GPR

The principle of GPR is based on the phenomenon that electromagnetic waves produce different reflections when they act on materials with different dielectric constants. Figure 1 shows the "scan" (a series of reflected signals detected) during tree roots detection. The GPR equipment moves in a preconfigured trajectory and emits electromagnetic waves to the ground, the electromagnetic waves are partially reflected at the roots and soil interface, and the rest of the electromagnetic waves continue to propagate downward until the signal is fully attenuated, as in Fig. 1a. After receiving the reflected electromagnetic waves from the roots, the receiving antenna records the electric field intensity change of the reflected waves in the time domain, forming an A-scan curve of the field intensity change with time as in Fig. 2a.

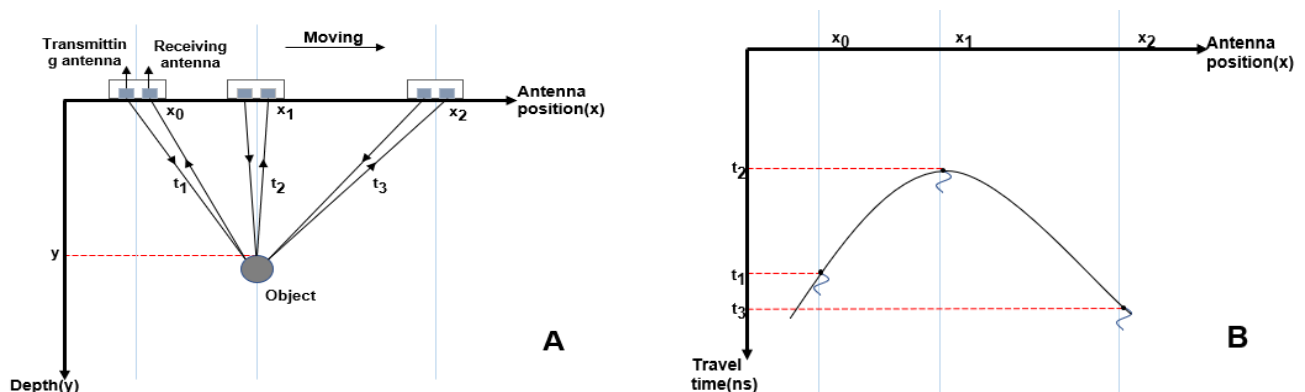


Fig. 1. Ground-penetrating radar object detection imaging principle graph (A) the radar signal is reflected by the buried object at positions (x_0 , x_1 , and x_2), and the reflection time (t_1 , t_2 , and t_3) is recorded and plotted below the radar. (B) Different A-scans form a reflection hyperbola during the movement.

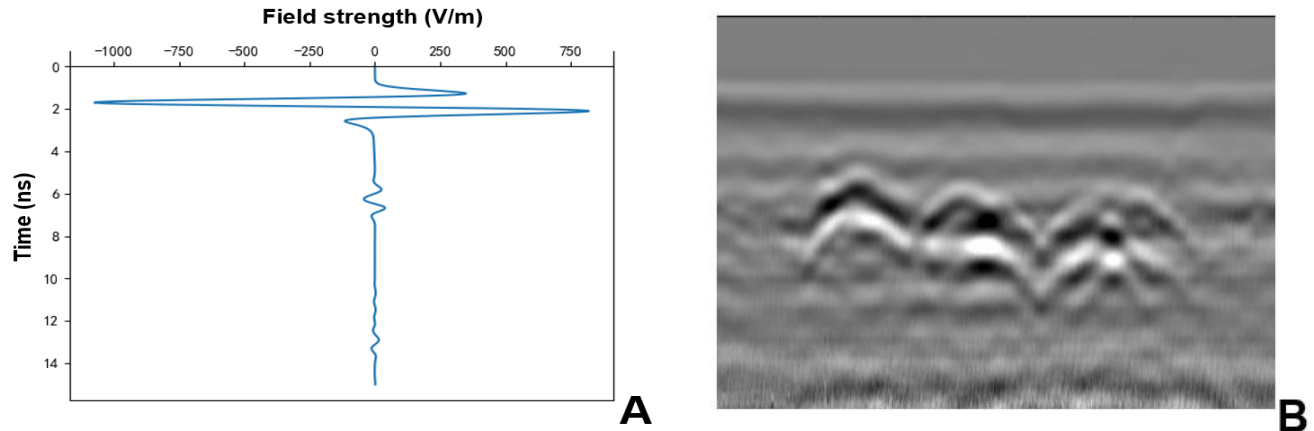


Fig. 2. A sample of real data: (A) an A-scan signal curve (B) a real B-scan image

When the transmitting and receiving antennas transmit electromagnetic waves to the ground once at x_0 of Fig. 1a, the A-scan plot of the x_0 in Fig. 1b is recorded by the receiving antenna. When the GPR moves backward in equal steps and transmits electromagnetic waves to the ground, a set of A-scan curves of the field intensity change caused by the reflected electromagnetic waves from the tree roots is recorded. Merging this set of A-scan curves to form the curve in Fig. 1b, which is the B-scan image, and a true B-scan image is displayed in Fig. 2b.

Image Acquisition

The trees detected in this paper were mainly distributed in Beijing, Shandong, Zhejiang, *etc.* The detection species mainly include willow, pine, cypress, *etc.* The measured trees were all isolated trees within a 5 m radius, ensuring that the detected roots were all associated with specific trees. The tree information was uploaded to the built website as shown in Fig. 3.

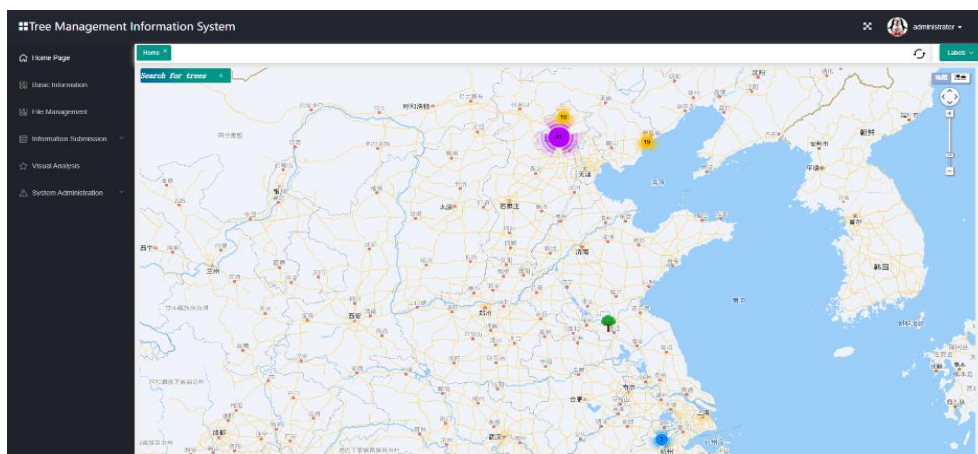


Fig. 3. All trees displayed on the map

The detection images of nearly one hundred trees' root systems were selected as the real dataset in the experiment. TRU tree radar, which is more compatible with the characteristics of tree roots, was chosen as the collection equipment (SIR3000T, GSSI,

USA). In practical applications, the antenna frequency of 900 MHz was selected, its maximal depth of detection was approximately 1 m, and the tracking interval and number of samples were 5 mm and 512, respectively. Meanwhile, the authors used a detection radius between 0.1 m and 3.8 m, detection depth of 0.6 m, and 0.75 m for on-site detection.

Dataset Construction

The composition of GPR image dataset consisted of four parts: acquisition of real data, simulated data generation, enhanced data generation, and data combination. The whole process is shown in Fig. 4.

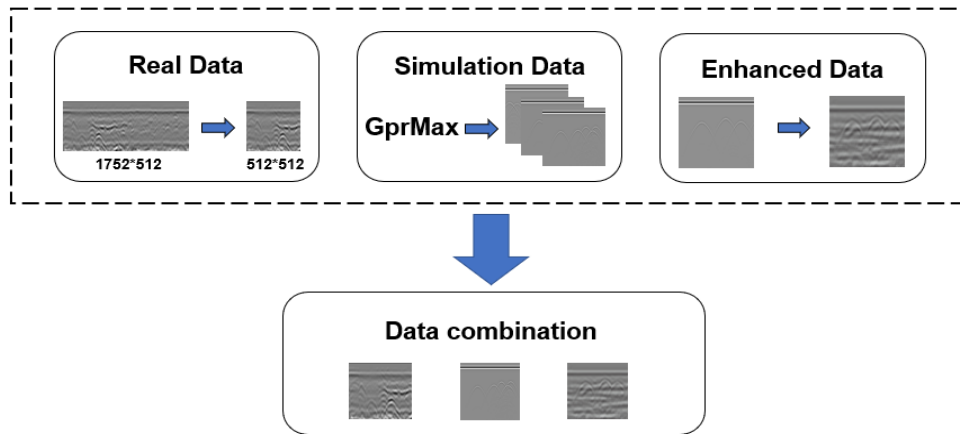


Fig. 4. The four parts of the dataset construction. The first section is the real data acquisition, which introduces the way to acquire GPR B-scan images. The second section is the generation of simulation data by GprMax. The third section is data augmentation. The fourth section utilizes the three kinds of data and constructs a hybrid dataset.

Acquisition of real data

To ensure the quality of the real dataset, the obtained real images were pre-processed. Nearly one thousand B-scan images of tree roots were screened, all with a height of 512 pixels and a width ranging from 148 to 3816 pixels. To ensure the consistency of data size, 336 high-quality data images were obtained by discarding images with widths less than 512 pixels and blurred images. Based on these images, the images with a width more than 512 pixels were cropped to obtain 759 B-scan images with a resolution of 512×512 pixels.

Simulated data generating

In this study, the simulation data had a crucial role in extending the diversity of the dataset. When generating simulated data, the parameter variables were controlled so that the GprMax software parameters remained consistent with the ground penetrating radar equipment parameters, where: the depth of the domain was 0.6 m, the lateral length was 6 m, the root had a radius from 0.01 to 0.035 m, the soil and root system have relative dielectric constants of 6 and 12 (Attia al Hagrey 2007; Liang *et al.* 2021), respectively, and the sampling number was 512, the antenna frequency of the GPR setting was 900 MHz, and a total of 759 simulated images were generated.

Enhanced data generation

The above analysis of the GPR target imaging process found that the tree roots showed hyperbolic structural features on the B-scan images, while different soil environments showed various background and noise features on the B-scan images. A GPR B-scan image was composed of three elements: hyperbolic structure, background, and noise features. When performing the conversion of a simulation image to a generated image, the complete hyperbolic structure features and nearly realistic background and noise features should be retained in the generated image. The tree environments in the experiment were different, which allowed the CycleGAN model to generate richer images, and the hyperbola, background, and noise features of the generated images were closer to the real images. Specifically, the hyperbolic feature information was fused with the background and noise for the real acquired GPR B-scan images, and it was difficult to distinguish the hyperbolic curve. Therefore, the hyperbolic image without background and noise was generated by style transformation, which could clearly show the hyperbolic structure and facilitate further study. At the same time, the generated images with different background and noise were generated through style transformation on the basis of retaining the simulation image hyperbolic features, increasing the diversity of the real dataset. The different backgrounds and noise were appended to the simulation images, making the generated images closer to the measured data.

The transformation of simulation images to generated images followed the CycleGAN architecture. The structure of this network consisted of two pairs of generators and discriminators. Generator A transformed the real image with features and generated the corresponding simulation image A'. The generated-simulation image A' was compared with the simulation image, expecting to obtain a generated-simulation image that could be faked as real, which could also be considered as a B-scan image with the background and noise removed. The specific architecture is shown in Fig. 5, which consists of two generators, GA and GB, and two discriminators, DA and DB.

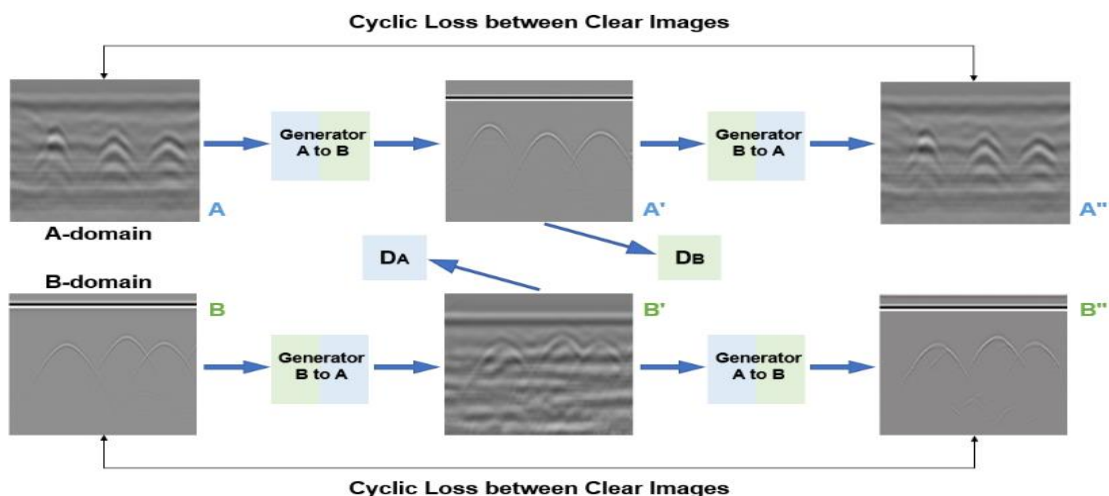


Fig. 5. Overall architecture of the proposed CycleGAN architecture for background and noise conversion of a single image

In particular, generator GA was used to generate B-domain style images from A-domain, and generator GB reverted the generated B-domain images to A-domain images. The discriminator DB was used to make the image generated by the generator GA as close

as possible to the B-domain style image, and the discriminator DA was used to make the image generated by the generator GB as similar as possible to the original B-domain original image, ensuring that when the image style was migrated, there is not only a change in style from the B-domain to the A-domain but also the features of the original B-domain original image still exist. The cyclic consistency loss allows us to train a model that does not require pairs of image instances with or without background and noise features.

The real dataset and the simulated dataset were used as A-domain and B-domain in the structure of CycleGAN, respectively. During training, the real image A in the A-domain was transformed into the corresponding generated image A' after the generator GA. A' had similar hyperbolic features as the real image. There was no noise and background, and the hyperbola was clearer, which could be used as the initial noise reduction of the real image, and A' was compared with the B-domain simulation image by DB to discriminate whether A' satisfies the conditions of the B-domain images. Similarly, the simulation image B in the B-domain was transformed into the corresponding generated image B'. B' retained the hyperbolic features of B and added noise and background. B' was compared with the real image in the A-domain by DA to discern whether B' satisfied the A-domain image. Meanwhile, A' was input to generator GB to generate the reduced image A'', and B' was input to generator GA to generate the reduced image B'', and the cyclic consistency loss of A'' with A and B'' with B was calculated so that the hyperbolic features of the original image were still retained while the background and noise of the generated image were changed.

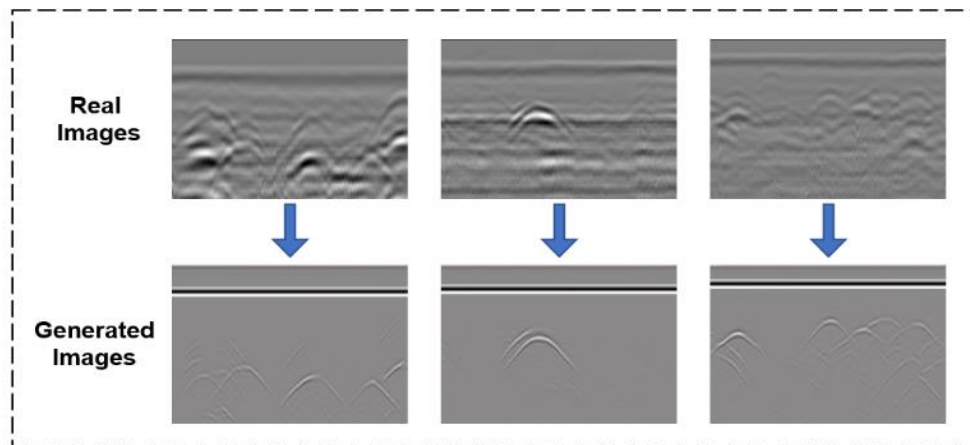


Fig. 6. Image style conversion diagram

The conversion of the real image to the simulation image is shown in Fig. 6, and it could be observed that the generated image retained most of the hyperbolic features in the real image. The converted image facilitated the next step of hyperbolic positioning and research. The disadvantage was that the generated images still had some missing features, and the hyperbolic features were more complex and less clear than the simulated images generated by GprMax. Such images were not convenient for hyperbolic labeling. Therefore, the real B-scan images, the generated B-scan images, and the simulation images generated by GprMax were used to build the hybrid dataset in the experiment. To validate the feasibility of generating the B-scan dataset. Both cosine similarity and SSIM were used to evaluate the similar relationship between the real image A and the restored image A''.

Structural Similarity Index measure (SSIM) is a single-scale image structural similarity, which is widely used as a measure of structural similarity between images. Its value is distributed in range of 0 to 1, and a value closer to 1 translates to more similarity between the two images, and a value closer to 0 means there is less similarity between the two images. The brightness, contrast, and structural similarity of the two images are reflected by the mean, standard deviation, and covariance of the images, which are calculated as follows.

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (1)$$

$$l(x, y) = (2u_x u_y + c_1) / (u_x^2 + u_y^2 + c_1) \quad (2)$$

$$c(x, y) = (2\sigma_x \sigma_y + c_2) / (\sigma_x^2 + \sigma_y^2 + c_2) \quad (3)$$

$$s(x, y) = \sigma_{xy} + c_3 / (\sigma_x \sigma_y + c_3) \quad (4)$$

where u_x and u_y indicate the average value of image x and y , respectively, σ_x and σ_y indicate the standard deviation of image x and y , respectively, and σ_{xy} indicates the covariance of image x and y . Then c_1 , c_2 , and c_3 are the very small numbers, which are designed to avoid the case of zero denominator in the above equation.

Cosine similarity is also called cosine distance. This method uses the cosine of the angle between two vectors in vector space to evaluate the difference between two images. The similarity (Eq. 5) is as follows,

$$Similarity(A, B) = A \cdot B / \|A\| \times \|B\| = \sum_{(i=1)}^n (A_i \times B_i) / \left(\sqrt{\sum_{(i=1)}^n A_i^2} \times \sqrt{\sum_{(i=1)}^n B_i^2} \right) \quad (5)$$

where A and B represent two vectors, the closer the angle is to 0° , the closer the cosine value is to 1, which means the more similar the between two vectors.

The authors calculated the cosine similarity between images in the A-domain and the corresponding reduced images A" to be 0.923, and the SSIM was 0.894. The results showed that the generated B-scan images retained the hyperbolic features of the real images and generated the background and noise features similar to the real images. Extending the real dataset was feasible by generating B-scan images with CycleGAN.

The 759 data of the simulation images were transformed by CycleGAN to obtain the generated image data, and original data of the three types of images were annotated using LabelImg, where YOLOv5 used TXT format, Faster R-CNN and YOLOv3 used XML format for annotation, and CenterNet used JSON format. The annotation information could indicate the location and size of the tree roots. Traditional offline data augmentation (random horizontal flip, random distortion, Gaussian blur, and random stretching) was performed on each type of labeled image, as shown in Table 1.

Table 1. Data Distribution

Image	Original Quantity	Enhanced Multipliers	Enhanced Quantity	Training Images	Validation Images	Testing Images
Real Images	759	15	12144	10000	1144	1000
Simulation Images	759	15	12144	10000	1144	1000
Generate Images	759	15	12144	10000	1144	1000

As shown in Fig. 7, the authors deployed the noise reduction function of CycleGAN to the designed tree management information system, and the developed tree management information system is able to process and analyze the GPR B-scan images online, which reduces the user's operation difficulty and improves the work efficiency.

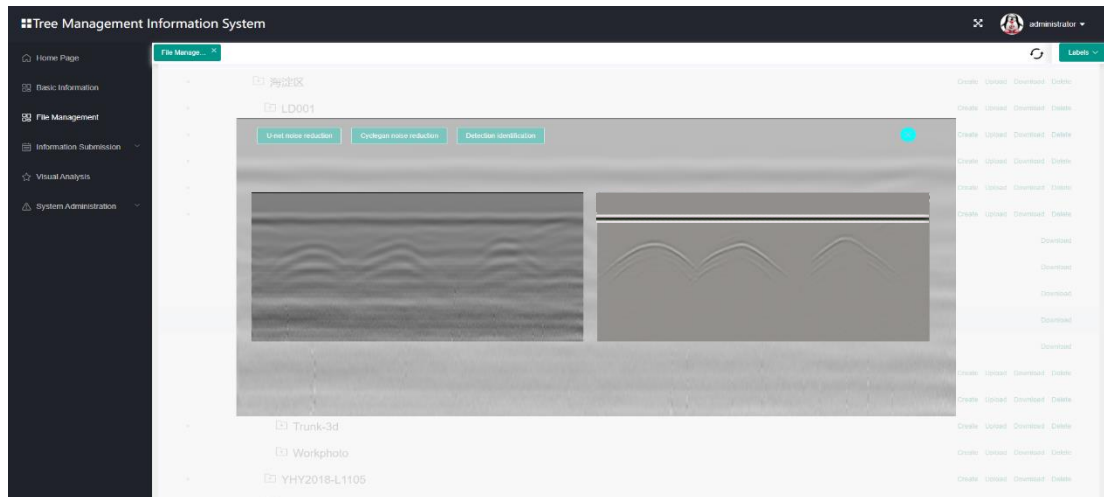


Fig. 7. Deployment of CycleGAN to a designed tree management information system

Data combination

After the above steps are completed, the combination of datasets is performed, a total of seven datasets are constituted, and the composition of the datasets is shown in Table 2.

Table 2. Composition of the Datasets

Dataset	Training Dataset	Validation Dataset	Testing Dataset
Real (R)	10000	1000	1200
Simulation (S)	10000	1000	
Generation (G)	10000	1000	
Real Generation (RG)	10000	1000	
Real Simulation (RS)	10000	1000	
Generation Simulation (GS)	10000	1000	
Real Generation Simulation (RGS)	10000	1000	

Except for the RGS training dataset consisting of 4000 real images, 3000 simulation images, and 3000 generated images, the remaining training datasets consisting of two types of data are each taken the first 5000 images of each data to form the final training datasets. To measure the training effects of all training datasets, 400 images from each of the three types of data are taken to form the final testing dataset, and all training models are tested with it.

Hyperbolic Detection Model

YOLOv5 network architecture

With the development of target detection technology, the YOLO series has been pursuing the best balance of speed and accuracy in real-time detection applications. YOLOv5, the latest achievement in recent years, has dramatically improved speed and accuracy compared to the previous series and has been applied to various fields with better results. YOLOv5 consists of Backbone (CSPDarknet), Neck (PANet), and Head (YOLO layer) parts as shown in Fig. 8.

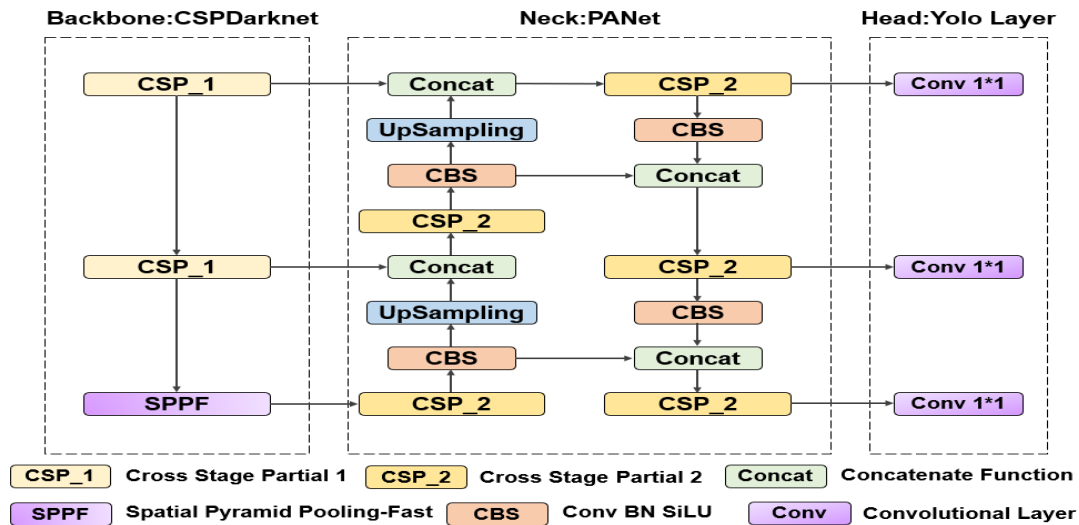


Fig. 8. YOLOv5 architecture

CSPDarknet reduces the parameters and computation of the model and the model's size, ensuring the operation speed and accuracy. The SPPF (Spatial Pyramid Pooling Fast) network is used to increase the receiver domain of the network, as shown in Fig. 9. The network uses multiple 5×5 -sized Maxpool layers to obtain richer features. Compared to the SPP (Spatial Pyramid Pooling), the SPPF serially passes the input through multiple Maxpool layers, obtaining the same computational results as the SPP but more efficiently. Then, using PANet (Path Aggregation Network) as a neck network can preserve the spatial information accurately. This network contributes to the correct positioning of pixels and forms a mask to better utilize the extracted features. When the image passes through each layer of the neural network, the feature complexity increases, while reducing the spatial resolution of the image. Thus, the pixel-level masks are not accurately recognized by the high-level features.

The FPN (Feature Pyramid Network) uses a top-down path to extract semantic-rich features and combines them with accurate location information. Meanwhile, CBL (Convolution, Batch Normalization, and leaky-ReLU) is replaced by CBS (Convolution, Batch Normalization, and SiLU), and the SiLU activation function has better nonlinear capabilities. The head part uses the head network of YOLOv3 to predict the obtained features.

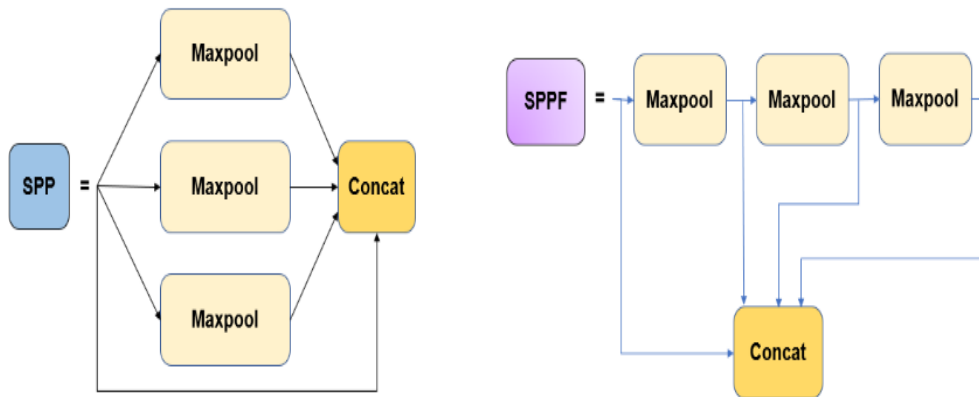


Fig. 9. Comparison between SPP and SPPF

Training setup

The training environment was on the basis of Python 3.8.6, PyTorch 1.7 (used in CenterNet, YOLOv3, and YOLOv5 models), and Python 3.6.13, tensorflow 1.11.0 (used in Faster R-CNN models). The computers used in all experiments were equipped with the following features: Intel(R) Core(TM) i5-9400 CPU, 16GB RAM, NVIDIA GeForce GTX 1660 Ti GPU, and a SAMSUNG 250G SSD hard drive.

This paper compares four different models, YOLOv3, YOLOv5, Faster R-CNN, and CenterNet. YOLOv5 has four models (S, M, L, and X) with different depths, and the S model was selected for training. The suitable model parameters were selected after a comprehensive consideration of the dataset and hardware. The hyperparameters of the model were set as follows: batch size was 16; momentum decay and weight decay were 0.8 and 0.0005, respectively; input size was 256×256 ; initial learning rate was 0.002; the epoch of YOLOv3, YOLOv5, and CenterNet models was 100; iteration of Faster R-CNN model was 10,000 iterations, which was approximately equal to 325 epochs, satisfying the basic training requirements; other default values were used.

Model Evaluation Indicators

In this paper, several quantifiable metrics were employed to evaluate the performance of the selected model quantitatively, including mean precision (mAP), precision (P), recall (R), and F1 score.

The precision and recall rate

In the object detection model, precision and recall are the two most basic evaluation indicators. Precision is defined as the percentage of all detected objects that are correctly detected, while recall is defined as the percentage of all detected positive samples that are correctly detected. The equations for these two metrics are as follows,

$$P = TP / (FP + TP) \quad (6)$$

$$R = TP / (FN + TP) \quad (7)$$

where TP is the number of correctly detected hyperbolas, FP is the number of non-hyperbolas treated as hyperbolas, and FN is the number of hyperbolic selections treated as non-hyperbolas.

The mAP and F1 score

The mean average precision is a composite metric that combines precision and recall. It is the average value of the average precision (AP) of all categories. In this study, mAP is equivalent to AP because only one object (hyperbola) is available. The *mAP* can be expressed as the area enclosed by the accuracy and recall curves, as in Eq. 8:

$$mAP = \int_0^1 P(R)dR \quad (8)$$

The *F1* score is used to assess the overall performance of the model. The calculation formula is shown in Eq. 9:

$$F1 = 2P \cdot R / P + R \quad (9)$$

RESULTS AND DISCUSSION

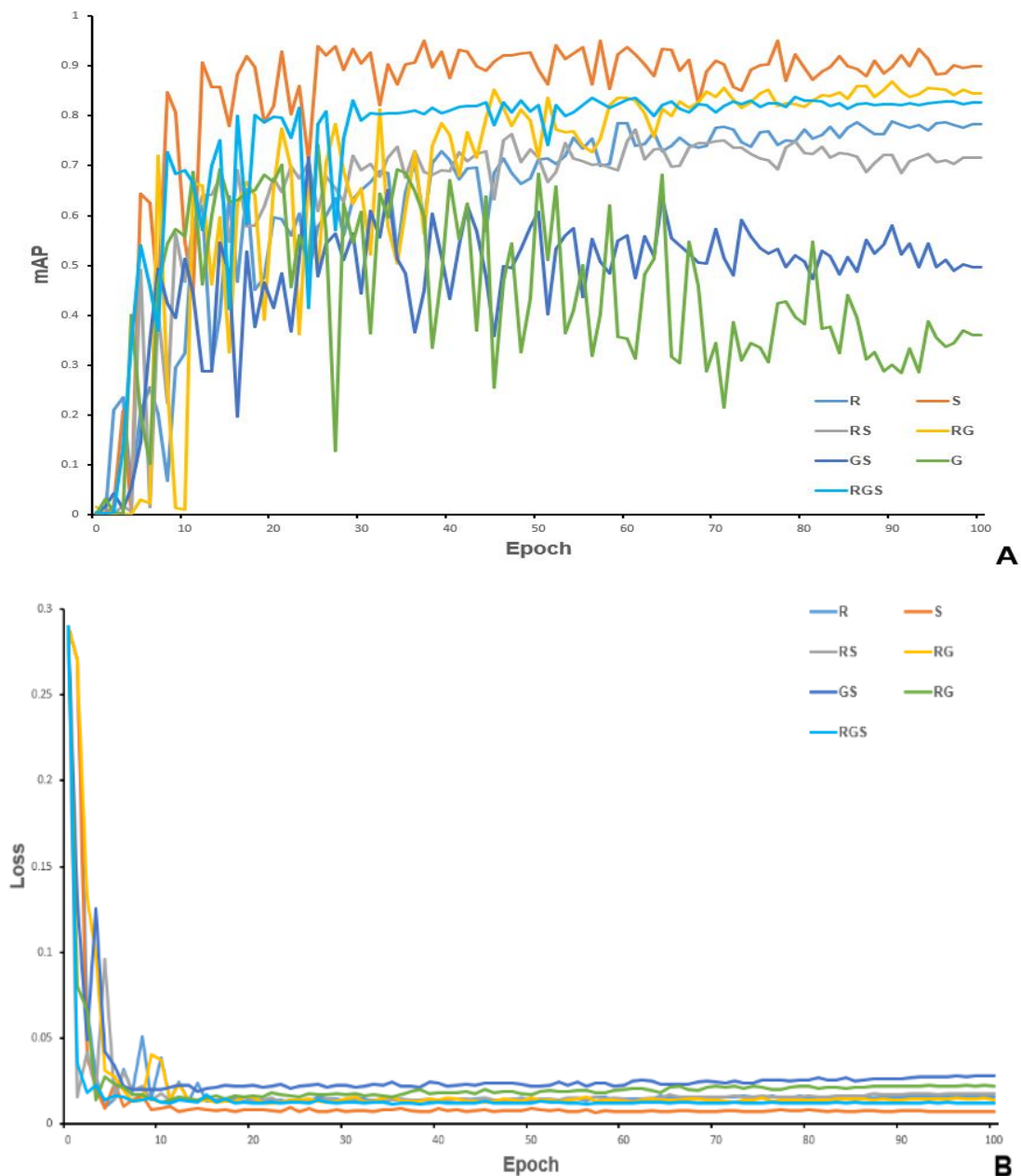
Analysis of Training Results

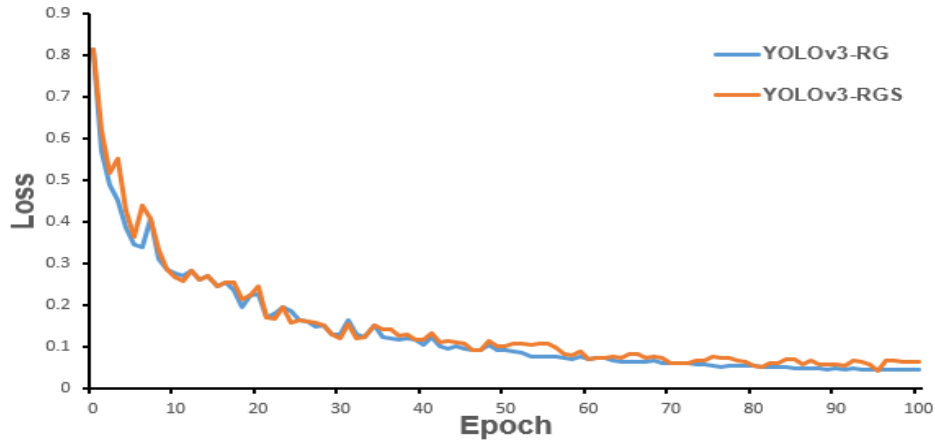
The value of loss indicates the difference between the predicted value and the true value. A low value of loss corresponds to a well-trained effect. At the same time, a higher mAP value also indicates that the trained model has a better performance. The loss curve and mAP curve of the model were compared, as shown in Fig. 10.

In Fig. 10(a,b), training on the use of YOLOV5 for the selected seven datasets, the simulated dataset had the lowest loss and the highest mAP, which was due to the absence of a noisy background from the simulation image, making it easier to detect and identify. However, the worst results were obtained when detecting real images, and it was almost impossible to recognize hyperbolas. Compared with the R and RS training models, the RG training model and the RGS training model had worse convergence loss and mAP after training. Both the G and GS training models without adding real B-scan images had poorer loss and mAP, and mAP was 49.25% and 33.78% lower for the G training model compared to the RG training model and the GS training model compared to the RGS training model, respectively. Therefore, the involvement of real B-scan images is necessary to obtain better training results. Comparing the results of the R training model and the RS training model, as well as the RG training model and the RGS training model, it can be seen that the mAP decreased 1.76% to 6.75% after adding the simulated data to the corresponding real B-scan dataset. This is because the number of dataset is sufficient to achieve good results when using R or RG dataset for training. If simulated data are added, the number of real data is reduced, while simulated data will disturb the judgment of the real data and affect the training effect. The better training effect of the RG dataset than the R dataset indicates that the addition of the generated data helps to increase the convergence effect of the model and improve the recognition ability of the model, mAP is improved 7.14%.

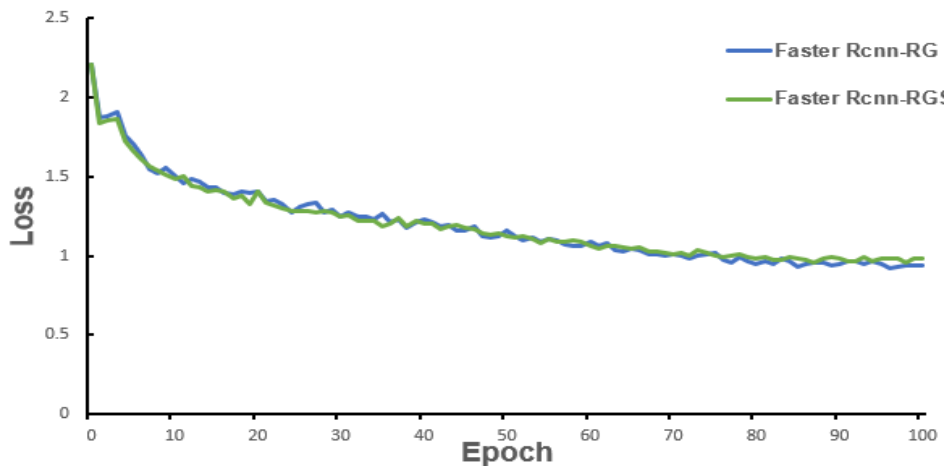
Due to the small amount of real data, the traditional way of training using the RS dataset keeps its mAP value at approximately 71.53%. Adding generated data to the RS dataset and training using the RGS dataset keeps its mAP at approximately 82.78%, with an 11.25% increase in mAP, which is noticeable. The results show that the generated data can both expand the real B-scan dataset and increase the performance and detection accuracy of the trained model. To further validate the effect of the generated data on the training effect, YOLOv3, Faster R-CNN, and CenterNet networks are used to train the RG dataset and RGS dataset.

Figure 10(c,d,e) corresponds to the Loss curves of the three networks after training. The Loss values of three networks after training were lower than 1.5, and the loss value of the RG model was lower than that of the RGS model, verifying that the simulation images increase the loss value of its model and decrease its mAP value. The mAP curves after training using different deep learning methods for the RG and RGS datasets are shown in Fig. 10f. The YOLOv5 model outperforms the YOLOv3, Faster R-CNN, and CenterNet models in terms of mAP during training (Fig. 10f). The mAP of the YOLOv5 model exceeds 80% on both datasets. The mAP values of both Faster R-CNN and YOLOv3 remain between 70% and 80%, and the mAP value of Faster R-CNN is slightly higher than that of YOLOv3. CenterNet has the worst training results, with mAP below 70%. Combining the loss and mAP values yields that YOLOv5 model has an impressive training performance.

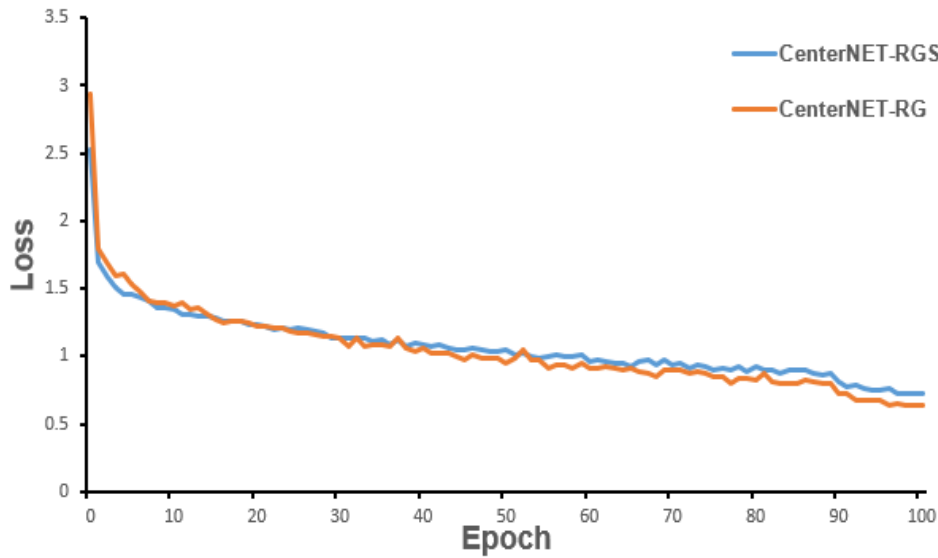




C



D



E

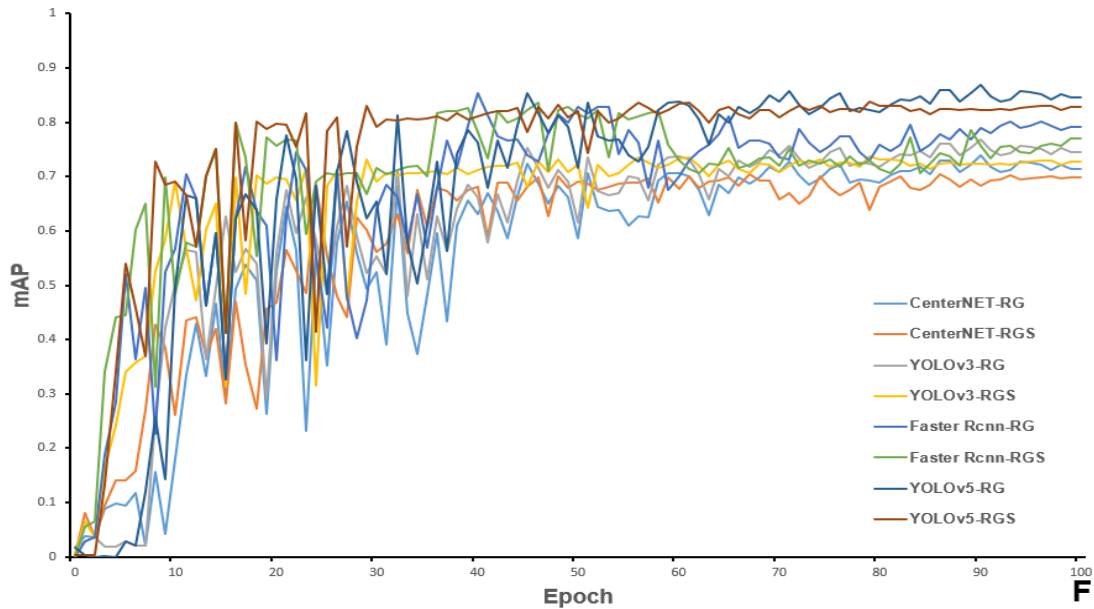


Fig. 10. Training results of the model: (A) The mAP values of the training process using YOLOv5 for 7 datasets. (B) Training Loss values for 7 datasets. (C) Loss curves of the YOLOv3 model. (D) Loss curves of the Faster R-CNN model. (E) Loss curves of the CenterNet model. (F) The mAP values of the different methods on each of the two datasets.

Table 3 summarizes all training results for the four networks corresponding to the eight models. The values of the evaluation indicators of the YOLOv5 model exceed the other models by 5% to 15% in terms of F1 scores and mAP values.

Table 3. Training Results of the Eight Models

Model \ Dataset	RG				RGS			
	P	R	mAP	F1	P	R	mAP	F1
YOLOv3	0.76	0.72	0.74	0.77	0.77	0.68	0.72	0.74
YOLOv5	0.84	0.81	0.85	0.83	0.88	0.75	0.82	0.81
Faster R-CNN	0.79	0.78	0.79	0.81	0.81	0.73	0.77	0.76
CenterNet	0.70	0.71	0.69	0.71	0.73	0.67	0.68	0.70

Analysis of Testing and Test Results

In this paper, all dataset models were tested with the same testset, and the test results are shown in Fig. 11. The RGS training model testing results are the best, with better detection for real data, simulated data, and generated data. The mAP value is improved approximately 10% compared to the RS training model, while the RG training model outperforms the R training model in terms of recall and mAP, confirming the reliability of the training results. It was shown that the generated data may improve the comprehensive performance of the training model.

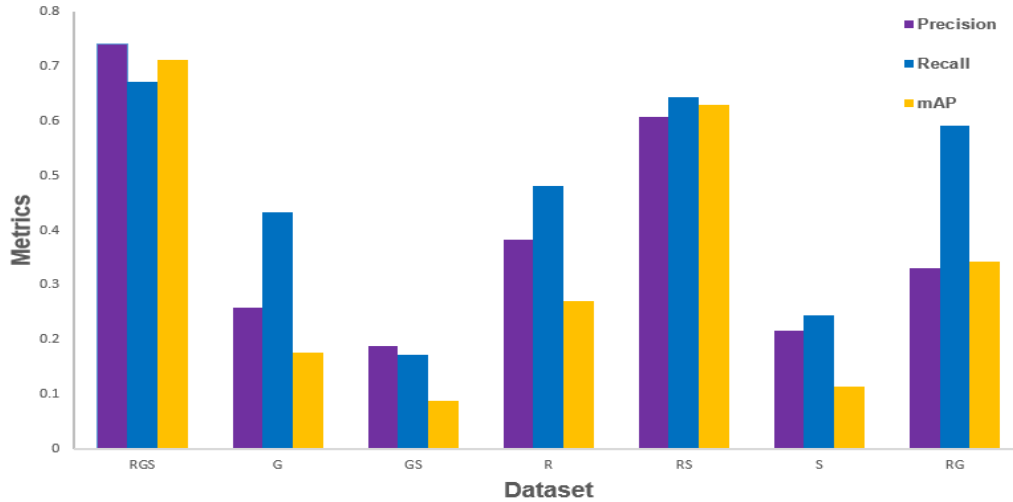


Fig. 11. Effect of testing set on seven dataset models(RGS: Real Generation Simulation;G: Generation;GS: Generation Simulation;R: Real;RS: Real Simulation;S: Simulation;RG: Real Generation)

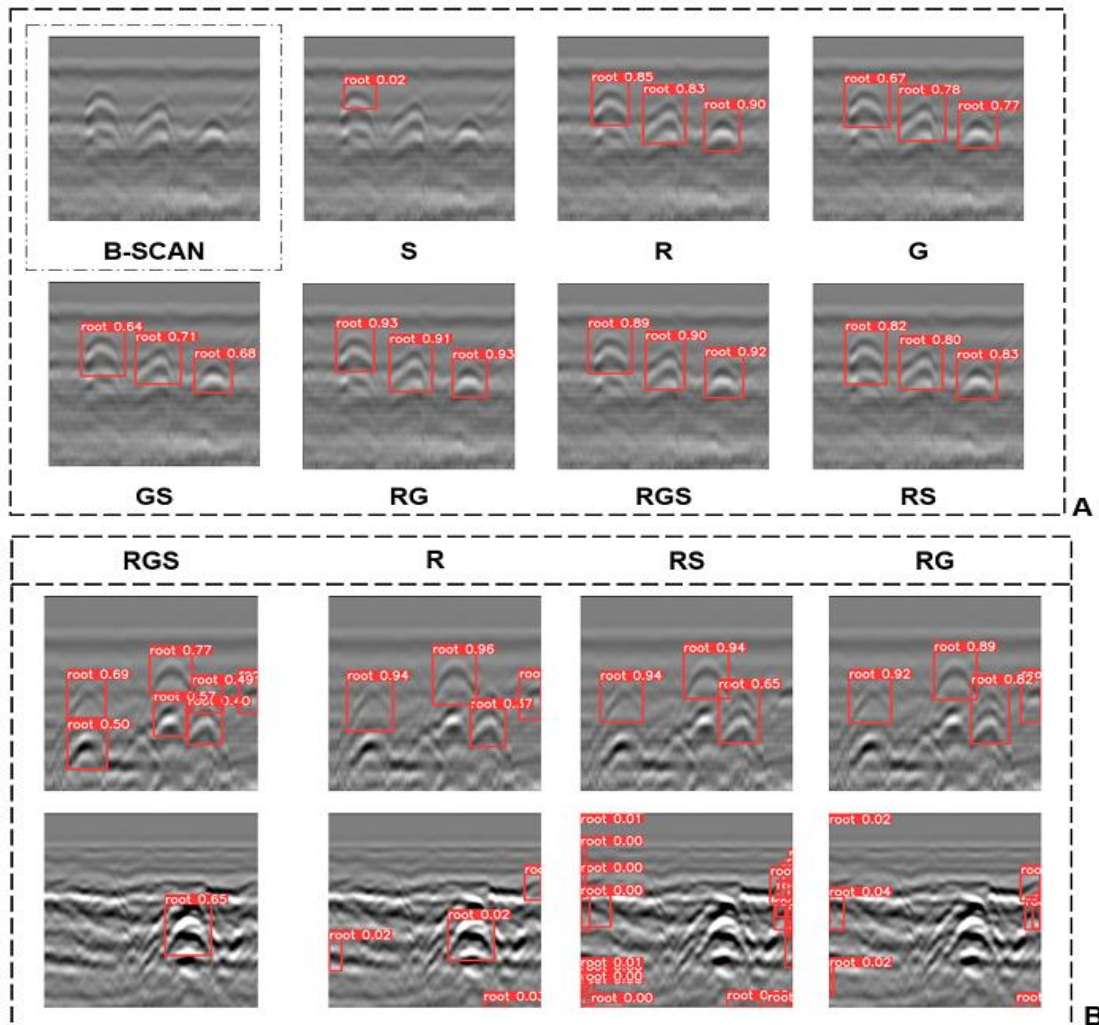


Fig. 12. Testing results of the data set: (A) Detection results of seven dataset models using YOLOv5; (B) Detection results of images under complex conditions

Figure 12a shows the partial detection results of YOLOv5 training seven dataset models on real B-scan images. There are three hyperbolic features in the real B-scan image, except for the S dataset training model, which is not detected completely, the rest of the dataset training models are detected completely, and the RG training model has the lowest confidence in detection with the mean value of only 0.68. The training model for the G dataset improved slightly with a confidence mean value of 0.74. The confidence mean values for the RS, R, RG, and the RGS training models were above 0.8, and the RG training models had the highest confidence mean value of 0.92.

To further verify the effect of various datasets, images with more complex backgrounds and hyperbolas with crossover cases for recognition were selected, with recognition confidence thresholds of 0.2 and 0.01, respectively, and the recognition results are shown in Fig. 12b. The recognition results show that the RGS training model had the best robustness and good recognition effect for various pairs of hyperbolic cases with a high recall rate. Especially for the hyperbolic images with a complex background, the recognition confidence of the RGS training model still reached 0.65, while the other dataset models can barely recognize the hyperbolas with complex backgrounds. The R, RS, and RG training models have high confidence in the recognition of identifiable curves when recognizing the presence of crossed hyperbolas, maintaining approximately 0.9 with a high accuracy rate, but half of the hyperbolas were not identified and the recall rate was low.

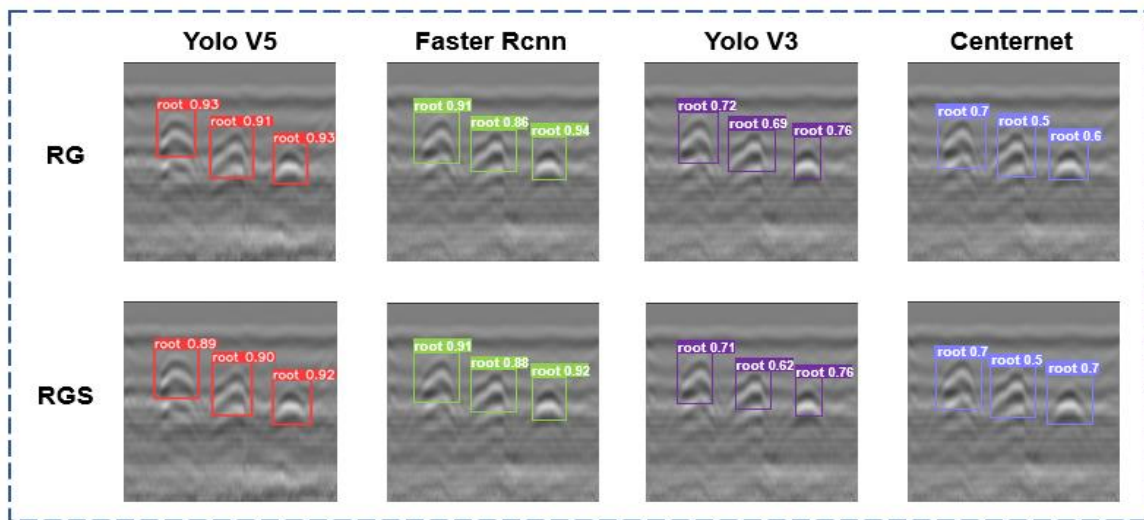


Fig. 13. Detection results of four methods for two datasets

Figure 13 shows the detection results of four deep learning methods for the RG dataset and the RGS dataset, all eight models identified and localized the three hyperbolas, but there were large differences in confidence levels. The CenterNet model had a minimum confidence value of 0.5, while in the Faster R-CNN, the maximum value was 0.94. The YOLOv3 confidence value ranged from 0.62 to 0.76. The YOLOv5 and the Faster R-CNN had similar results, with the mean confidence values reaching 0.9 or higher, and YOLOv5 was slightly higher than Faster R-CNN. The above results show that YOLOv5 model detection recognition performance using the enhanced dataset is the best.

CONCLUSIONS

1. Based on the deep learning method to achieve automatic recognition and localization of hyperbolas, the YOLOv5 model had the best detection results compared to YOLOv3, Faster R-CNN, and CenterNet. The average accuracy of the models all amounted more than 80%, and the recognition rate and recall rate of the YOLOv5 training model reached 84.45% and 81.23%, respectively, and the target hyperbolas had strong detection for different detection tasks and recognition capability for different detection tasks.
2. The real dataset was enhanced by CycleGAN, and the enhanced image had high similarity with the real images. In addition, the hyperbolic features were kept intact. At the same time, the background and noise of the real images could be removed. The removed image was similar to the simulation image, and only a few hyperbolic features were lost, resulting in a good overall effect for the next step of the study.
3. Compared with the traditional RS dataset, the RGS dataset and the RG dataset train the model better, the loss and mAP of the training process were higher than in other datasets, and the actual detection effect was the best, with the highest confidence of detecting hyperbola.
4. When constructing the GPR B-scan training dataset, there must be a large amount of real data, and less real data would lead to poor actual results; generating B-scan data had similar background and noise and hyperbolic features as the real data, and the expansion of the real dataset, while being able to improve the effectiveness of the training model, the mAP improvement reached more than 3%.
5. The extracted method was applied to the web-based management information system, simplifying the operation process and making the processing results easy to observe.
6. By the method in this paper, the user can clearly observe the number of roots, root density, and root distribution, which is useful for predicting root diameter and other information. In the future, such information can be used for assessment of tree growth and health.

ACKNOWLEDGEMENTS

This study was supported by the National Natural Science Foundation of China (Grant No. 32071679) and the Beijing Municipal Natural Science Foundation (Grant No. 6202023).

REFERENCES CITED

- Aboudourib, A., Serhir, M., and Lesselier, D. (2021). "A processing framework for tree-root reconstruction using ground-penetrating radar under heterogeneous soil conditions," *IEEE Transactions on Geoscience and Remote Sensing* 59(1), 208-219. DOI: 10.1109/TGRS.2020.2993719

- Alani, A. M., Ciampoli, L. B., Lantini, L., Tosti, F., and Benedetto, A. (2018). "Mapping the root system of matured trees using ground penetrating radar," in: *17th International Conference on Ground Penetrating Radar (GPR)*, IEEE, Rapperswil, Switzerland, pp. 1-6. DOI: 10.1109/ICGPR.2018.8441535
- Attia al Hagrey, S. (2007). "Geophysical imaging of root-zone, trunk, and moisture heterogeneity," *Journal of Experimental Botany* 58(4), 839-854. DOI: 10.1093/jxb/erl237
- Danjon, F., and Reubens, B. (2008). "Assessing and analyzing 3D architecture of woody root systems, a review of methods and applications in tree and soil stability, resource acquisition and allocation," *Plant and Soil* 303(1–2), 1-34. DOI: 10.1007/s11104-007-9470-7
- Davletshin, A., Ko, L. T., Milliken, K., Periwal, P., Wang, C.-C., and Song, W. (2021). "Detection of framboidal pyrite size distributions using convolutional neural networks," *Marine and Petroleum Geology* 132, Article ID 105159. DOI: 10.1016/j.marpetgeo.2021.105159
- Dewantara, D., and Parnadi, W. W. (2022). "Automatic hyperbola detection and apex extraction using convolutional neural network on GPR data," *Journal of Physics: Conference Series* 2243(1), article ID 012027. DOI: 10.1088/1742-6596/2243/1/012027
- Gill, R. A., and Jackson, R. B. (2000). "Global patterns of root turnover for terrestrial ecosystems: RESEARCH Root turnover in terrestrial ecosystems," *New Phytologist* 147(1), 13-31. DOI: 10.1046/j.1469-8137.2000.00681.x
- Guo, L., Chen, J., Cui, X., Fan, B., and Lin, H. (2013). "Application of ground penetrating radar for coarse root detection and quantification: A review," *Plant and Soil* 362(1–2), 1-23. DOI: 10.1007/s11104-012-1455-5
- Hou, F., Lei, W., Li, S., and Xi, J. (2021). "Deep learning-based subsurface target detection from GPR scans," *IEEE Sensors Journal* 21(6), 8161-8171. DOI: 10.1109/JSEN.2021.3050262
- Kemna, A., Vanderborght, J., Kulesa, B., and Vereecken, H. (2002). "Imaging and characterisation of subsurface solute transport using electrical resistivity tomography (ERT) and equivalent transport models," *Journal of Hydrology* 267(3–4), 125-146. DOI: 10.1016/S0022-1694(02)00145-2
- LaBrecque, D. J., and Yang, X. (2001). "Difference inversion of ERT data: A fast inversion method for 3-D in situ monitoring," *Journal of Environmental and Engineering Geophysics* 6(2), 83-89. DOI: 10.4133/JEEG6.2.83
- Li, S., Fu, S., and Zheng, D. (2022). "Rural built-up area extraction from remote sensing images using spectral residual methods with embedded deep neural network," *Sustainability* 14(3), article ID 1272. DOI: 10.3390/su14031272
- Liang, H., Fan, G., Li, Y., and Zhao, Y. (2021). "Theoretical development of plant root diameter estimation based on gprMax data and neural network modelling," *Forests* 12(5), article no. 615. DOI: 10.3390/f12050615
- Liu, G., Nouaze, J. C., Touko Mbouembe, P. L., and Kim, J. H. (2020). "Yolo-tomato: A robust algorithm for tomato detection based on yolov3," *Sensors* 20(7), article no. 2145. DOI: 10.3390/s20072145
- Ma, H., Liu, Y., Ren, Y., and Yu, J. (2019). "Detection of collapsed buildings in post-earthquake remote sensing images based on the improved yolov3," *Remote Sensing* 12(1), article no. 44. DOI: 10.3390/rs12010044

- Mihai, A. E., Gereaa, A. G., Curioni, G., Atkins, P., and Hayati, F. (2019). "Direct measurements of tree root relative permittivity for the aid of GPR forward models and site surveys," *Near Surface Geophysics* 17(3), 299-310. DOI: 10.1002/nsg.12043
- Pettinelli, E., Di Matteo, A., Beaubien, S. E., Mattei, E., Lauro, S. E., Galli, A., and Vannaroni, G. (2014). "A controlled experiment to investigate the correlation between early-time signal attributes of ground-coupled radar and soil dielectric properties," *Journal of Applied Geophysics* 101, 68-76. DOI: 10.1016/j.jappgeo.2013.11.012
- Ren, S., He, K., Girshick, R., and Sun, J. (2017). "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39(6), 1137-1149. DOI: 10.1109/TPAMI.2016.2577031
- Reubens, B., Poesen, J., Danjon, F., Geudens, G., and Muys, B. (2007). "The role of fine and coarse roots in shallow slope stability and soil erosion control with a focus on root system architecture: A review," *Trees* 21(4), 385-402. DOI: 10.1007/s00468-007-0132-4
- Riedell, W. E., and Osborne, S. L. (2017). "Monolith root sampling elucidates western corn rootworm larval feeding injury in maize," *Crop Science* 57(6), 3170-3178. DOI: 10.2135/cropsci2017.04.0218
- Rosati, R., Romeo, L., Silvestri, S., Marcheggiani, F., Tiano, L., and Frontoni, E. (2020). "Faster r-cnn approach for detection and quantification of DNA damage in comet assay images," *Computers in Biology and Medicine* 123, article ID 103912. DOI: 10.1016/j.combiomed.2020.103912
- Sarro, W. S., Assis, G. M., and Ferreira, G. C. S. (2021). "Experimental investigation of the UPV wavelength in compacted soil," *Construction and Building Materials* 272, article ID 121834. DOI: 10.1016/j.conbuildmat.2020.121834
- Seyfried, D., and Schoebel, J. (2016). "Ground penetrating radar for asparagus detection," *Journal of Applied Geophysics*, 126, 191-197. DOI: 10.1016/j.jappgeo.2016.01.022
- Tanoli, W. A., Sharafat, A., Park, J., and Seo, J. W. (2019). "Damage prevention for underground utilities using machine guidance," *Automation in Construction* 107, article ID 102893. DOI: 10.1016/j.autcon.2019.102893
- Thanh Le, V. N., Truong, G., and Alameh, K. (2021). "Detecting weeds from crops under complex field environments based on Faster rcnn," in: *2020 IEEE Eighth International Conference on Communications and Electronics (ICCE)*, IEEE, Phu Quoc Island, Vietnam, pp. 350-355. DOI: 10.1109/ICCE48956.2021.9352073
- Todkar, S. S., Baltazart, V., Ihamouten, A., Dérobert, X., and Guilbert, D. (2021). "One-class SVM based outlier detection strategy to detect thin interlayer debondings within pavement structures using Ground Penetrating Radar data," *Journal of Applied Geophysics* 192, article ID 104392. DOI: 10.1016/j.jappgeo.2021.104392
- Wang, Y., and Li, X. (2015). "Experimental study on cracking damage characteristics of a soil and rock mixture by UPV testing," *Bulletin of Engineering Geology and the Environment* 74(3), 775-788. DOI: 10.1007/s10064-014-0673-x
- Wen, J., Li, Z., and Xiao, J. (2020). "Noise removal in tree radar b-scan images based on shearlet," *Wood Research* 65(1), 1-12. DOI: 10.37763/wr.1336-4561/65.1.001012
- Xiang, Z., Rashidi, A., and Ou, G. (2019). "An improved convolutional neural network system for automatically detecting rebar in GPR data," in: *Computing in Civil Engineering 2019*, American Society of Civil Engineers, Atlanta, GA, USA, pp. 422-429. DOI: 10.1061/9780784482438.054

- Yao, S., Chen, Y., Tian, X., Jiang, R., and Ma, S. (2020). “An improved algorithm for detecting pneumonia based on yolov3,” *Applied Sciences* 10(5), article no. 1818. DOI: 10.3390/app10051818
- Yarak, K., Witayangkurn, A., Kritiyutanont, K., Arunplod, C., and Shibasaki, R. (2021). “Oil palm tree detection and health classification on high-resolution imagery using deep learning,” *Agriculture* 11(2), article no. 183. DOI: 10.3390/agriculture11020183
- Zhang, X., Xue, F., Wang, Z., Wen, J., Guan, C., Wang, F., Han, L., and Ying, N. (2021). “A novel method of hyperbola recognition in ground penetrating radar (GPR) b-scan image for tree roots detection,” *Forests* 12(8), article no. 1019. DOI: 10.3390/f12081019
- Zhou, X., Chen, H., and Hao, T. (2019). “Efficient detection of buried plastic pipes by combining GPR and electric field methods,” *IEEE Transactions on Geoscience and Remote Sensing* 57(6), 3967-3979. DOI: 10.1109/TGRS.2018.2889248
- Zheng, J., Li, W., Xia, M., Dong, R., Fu, H., and Yuan, S. (2019). “Large-scale oil palm tree detection from high-resolution remote sensing images using Faster-rcnn,” in: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE, Yokohama, Japan, pp. 1422-1425. DOI: 10.1109/IGARSS.2019.8898360
- Zhou, J., Tian, Y., Yuan, C., Yin, K., Yang, G., and Wen, M. (2019). “Improved UAV opium poppy detection using an updated yolov3 model,” *Sensors* 19(22), article no. 4851. DOI: 10.3390/s19224851
- Zhu, J.-Y., Park, T., Isola, P., and Efros, A. A. (2017). “Unpaired image-to-image translation using cycle-consistent adversarial networks,” in: *2017 IEEE International Conference on Computer Vision (ICCV)*, IEEE, Venice, Italy, pp. 2242-2251. DOI: 10.1109/ICCV.2017.244

Article submitted: August 11, 2022; Peer review completed: November 5, 2022; Revised version received and accepted: November 8, 2022; Published: November 16, 2022.
DOI: 10.15376/biores.18.1.484-504