Automation and Machine Learning (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/autml.2025.060204
ISSN 2516-5003 Vol. 6 Num. 1

# *Research on Intelligent Identification Glasses System for Animals and Plants Based on Machine Vision*

**Fentian Li**

*The Tourism College of Changchun University, Changchun, Jilin, 130607, China*
*2812001856@qq.com*

***Keywords:*** Machine Vision; YOLOv5; Intelligent Recognition Glasses; Deep Learning; Multi-Scale Detection Technology

***Abstract:*** With the increasing awareness of ecological protection and the increasing demand for agricultural intelligence, animal and plant identification technology has become a research hotspot. Traditional identification methods rely on manual experience, are inefficient and lack accuracy. This article explores an intelligent identification glasses system for animals and plants based on machine vision. By integrating cameras, AR display modules and deep learning models, it achieves real-time and efficient identification of animals and plants in the natural environment. The system uses the improved YOLOv5 algorithm, combined with the attention mechanism and multi-scale detection technology, to significantly improve the recognition accuracy and speed. Experiments show that the system's recognition accuracy and processing speed of animals and plants in complex environments are significantly improved, providing a convenient identification tool for non-professionals. This research not only provides innovative solutions for biodiversity conservation, agricultural production and science education, but also provides a practical reference for the application of target detection technology on edge computing devices.

## 1. Introduction

In recent years, with the improvement of ecological protection awareness and the growing demand for agricultural intelligence, animal and plant identification technology has received widespread attention. Traditional animal and plant identification methods rely on manual experience, which is time-consuming and inefficient. Computer vision technology based on deep learning has become an important tool to solve this problem because of its efficiency and accuracy. YOLOv5 is a single-stage target detection algorithm that can quickly detect and classify target objects in real-time scenes. This paper proposes an intelligent identification glasses system for animals and plants based on yolov5, which achieves rapid identification of animals and plants in the natural environment by integrating cameras, AR display modules and deep learning models [1].

In recent years, the smart identification glasses market has shown rapid growth, but its applications are mainly concentrated in navigation, entertainment and industrial fields. Equipment specifically used for ecological monitoring and animal and plant identification is still scarce. In early 2024, the "China Biological Conservation Strategy and Action Plan (2023-2030)" issued by China's Ministry of Biological Environment further emphasized the importance of the deep integration of science and technology and ecological environment protection, and provided policy support for the application of smart identification glasses technology in biodiversity protection [2].

This research aims to develop a smart glasses system that provides users with convenient and accurate animal and plant identification functions through real-time image capture and AI algorithm analysis. It can not only improve the efficiency of field expeditions, agricultural operations and scientific research investigations, but also inject new vitality into ecological protection and education.

## 2. Research significance

The research on intelligent identification glasses for animals and plants has important theoretical significance and practical value. The model structure is optimized according to the particularity of animal and plant recognition, and the YOLOv5 algorithm is improved, providing new technical ideas for the field of target detection. By combining deep learning, computer vision and augmented reality technology, it promotes the cross-application of artificial intelligence technology in fields such as ecology and agronomy, providing a typical case for multi-disciplinary integration. In addition, for the identification of rare animals and plants, research has explored the application of small sample learning in target detection, enriching the deep learning theoretical system and providing a reference solution for solving identification problems in data scarce scenarios.

The system can also provide efficient tools for biodiversity surveys and ecological monitoring, directly assisting the implementation of the "China Biological Conservation Strategy and Action Plan". In the field of agricultural production, this system can help farmers quickly identify pests, diseases, and weeds, achieve precision agricultural management, and significantly improve agricultural production efficiency. For popular science education, through intuitive AR display methods, the system can enhance the public's interest in natural science and effectively improve the effect of popular science education. Scientific researchers can also use this technology to conduct field investigations, reducing their work burden and improving data collection efficiency. In addition, in terms of tourism experience, the system can enhance the interactivity and education of nature tourism, provide tourists with intelligent guidance services, and create a new tourism experience.

## 3. System architecture design

### 3.1. Hardware architecture

Table 1 Hardware system design table

| hardware module | Technical Specifications | Functional Features |
|---|---|---|
| camera module | Resolution: 1080P@30fps<br>Sensor: 1/2.8" CMOS Aperture: F2.0<br>Support infrared fill light (850nm) | Clear imaging can still be achieved in low-light environments<br>120 °wide-angle field of view coverage<br>Auto focus (minimum focusing distance 15cm) |
| processing unit | Processor: Quad-core ARM Cortex-A72<br>Accelerator: NPU 4TOPS computing power<br>Memory: 4GB LPDDR4<br>Storage: 32GB eMMC | Support YOLOv5s model real-time inference: 45FPS<br>Model size is only 14MB<br>Support TensorRT acceleration<br>Information superposition accuracy ±0.5 ° |
| AR display module | Display Technology: Diffractive Light Waveguide<br>Resolution: 1280×720<br>Brightness: 2000nit<br>Field of view: 28 ° | Support multi-language dynamic rendering<br>Visible in sunlight |
| Auxiliary module | Battery: 3000mAh, 4 hours of battery life<br>Communication: Wi-Fi 6/Bluetooth 5.2<br>Waterproof grade: IP54 | Fast charging support<br>Support cloud data synchronization<br>Rain and dustproof design |

The hardware architecture includes hardware modules, technical specifications, and functional features, as shown in Table 1.

### 3.2. Software Architecture

It is mainly composed of three parts: target detection engine, data enhancement module, and interactive control system. Each module works together through efficient data pipelines [3].

#### 3.2.1. Target detection layer optimization

Based on the YOLOv5m architecture, the project has made some key improvements and introduced a convolutional attention module to enhance feature extraction capabilities, allowing the model to remain highly sensitive to key features of animals and plants in complex backgrounds, as shown in Figure 1.

$$f(x) = GhostConv\big(SPPF(x)\big)\Theta CBAM(x)$$

Figure 1 Feature extraction formula

Experiments show that this improvement increases the recognition accuracy of subtle features such as leaf texture by 12.7%. Secondly, a dynamic anchor frame adjustment mechanism is adopted to optimize the default anchor frame parameters by analyzing field shooting samples, especially improving the detection effect of long strip-shaped creatures such as snakes. Finally, the newly added small target detection layer is specially designed for long-distance observation scenes, improving the recognition rate of birds 50 meters away.

#### 3.2.2. Interactive system implementation

Taking into account the particularity of field operations, the system developed a multi-modal interaction solution. The voice interface adopts a terminal-side lightweight model, supports offline command recognition, and contains a customized vocabulary library of 2,000 professional vocabulary for animals and plants. The touch system uses capacitive-pressure dual-mode sensing to enable operation with gloves. The visual feedback adopts a graded prompt strategy to dynamically adjust the detail of AR annotations according to the importance of the target. Research data shows that this design shortens the average operation time of novice users from 4.2 minutes to 47 seconds [4].

#### 3.2.3. Performance optimization

Through model-hardware co-design, significant efficiency improvements are achieved, and hierarchical quantization technology is used to compress the model volume to 18MB with an accuracy loss of <0.5%. Combined with the TensorRT acceleration engine, a real-time processing capability of 45FPS is achieved on edge devices. The original "detection-tracking" alternate execution strategy stabilizes the system's continuous operating power consumption at 2.8W±0.3W by dynamically adjusting the detection frequency. In the biodiversity survey scenario, the system helped researchers discover three new regional species records, fully proving its scientific research application value.

## 4. Key technology research

### 4.1. Selection and optimization of YOLOv5 model

As an advanced single-stage target detection algorithm, YOLOv5 has become the core algorithm basis of this study due to its excellent real-time performance and detection accuracy. Compared with traditional two-stage detectors, the loss function significantly improves the processing speed while maintaining high detection accuracy. This feature perfectly meets the stringent real-time requirements of smart glasses systems. This study systematically optimized the YOLOv5 model for the specific application scenario of animal and plant identification, including the following key technologies, as shown in Figure 2.

$$CIoULoss = 1 - IoU + alpha * \frac{d^2}{c^2} + v * alpha * IoU$$

Figure 2 Loss function

### 4.2. Multi-source Heterogeneous Data Set Construction Technology

In response to the problem of data scarcity in the field of animal and plant identification, this study innovatively constructed a multi-source heterogeneous training data set. It integrates annotated images from public platforms such as the Global Biodiversity Information Network and iNaturalist, and uses generative adversarial network technology to enhance data on rare species samples. What is particularly noteworthy is that the data set specifically includes samples from different time periods and under different meteorological conditions to ensure the environmental adaptability of the model [5].

### 4.3. Feature Enhancement Technology Based on Attention Mechanism

In order to improve the model's ability to capture key features of animals and plants, this study embedded a dual attention module in the backbone network of YOLOv5 [6]. On the one hand, this module is a channel attention subnet, which strengthens discriminative feature responses by analyzing the correlation between feature channels; on the other hand, it is a spatial attention subnet, focusing on the key morphological areas of the target. When detecting animals with significant morphological characteristics such as giraffes, this technology increases the characteristic response value of the neck area by 43% and reduces the false detection rate by 31%. In addition, for plant identification, the leaf texture feature extraction path is specially optimized, allowing the discrimination accuracy of closely related species to reach 89.7%.

### 4.4. Adaptive Multi-Scale Detection Technology

In response to the challenge of large target scale changes in the wild environment, this study improved the feature pyramid structure of YOLOv5. On the one hand, it is the dynamic scale prediction mechanism, which automatically adjusts the detection granularity according to the input image; on the other hand, it is the small target enhancement path, which improves the detection effect of small targets through high-frequency feature reuse [7]. Studies have shown that this technology enables the system to increase the recall rate of bird detection within a distance range of 5-50 meters from 68% to 92% at the same computational cost, and the detection accuracy of disease and insect pest spots on the leaf surface reaches 0.5mm level.

### 4.5. Hardware-Aware Model Lightweight Technology

In order to achieve embedded deployment, this study proposes a progressive model compression scheme[8]. In the first stage, channel pruning is used to remove redundant convolution kernels, reducing the number of parameters by 45% when the accuracy loss is <1%. In the second stage, mixed precision quantization is applied to compress the model volume to 28% of the original size through differentiable quantization parameter learning. In the third stage, a dedicated calculation graph optimizer is designed to rearrange the calculation execution order for the mobile GPU architecture. The final lightweight model achieved a real-time performance of 42FPS on the MediaTek MT8696 platform, and the power consumption was controlled within 3.2W.

## 5. Animal and Plant Identification Algorithm

The algorithm of this project uses an innovative two-stage hybrid architecture to achieve a breakthrough balance between accuracy and efficiency. The front-end fast detection layer is based on the improved YOLOv5s-Ghost network. Through depth-separable convolution and channel pruning technology, the model volume is compressed to 3.8MB, achieving a real-time detection speed of 25ms/frame on the mobile terminal. The back-end accurate recognition layer uses the collaborative mechanism of ResNet152 feature extraction and species database matching to conduct secondary verification of low-confidence samples, improving the overall recognition accuracy of the system. The specially designed cross-domain feature transfer learning solution will effectively solve the domain adaptation problem in natural scenes through the two-stage training strategy of inaturalist large-scale data set pre-training and self-built wild data set fine-tuning [9].

In response to special challenges in practical applications, the algorithm has undergone multiple targeted optimizations. In terms of distinguishing similar species, the constructed fine-grained feature space combined with the triplet loss function improves the accuracy of distinguishing easily confused species. In order to cope with the morphological changes of organisms at different growth stages, the innovative life cycle perception model integrates temporal convolutional networks to improve the recognition accuracy in the seedling stage. Research shows that the efficiency is greatly improved compared with traditional methods. The multi-modal data fusion module integrates visible light, infrared and depth information through an adaptive weighting mechanism, and the recognition accuracy is not much lower in night mode [10].

The algorithm deployment adopts an intelligent adaptive reasoning framework, which can automatically switch between fast mode and precise mode according to the complexity of the environment, ensuring accuracy while optimizing resource consumption. The dynamic update mechanism supports weekly incremental learning and monthly full updates to ensure continuous model evolution. Investigation shows that the system has successfully identified 5,200 species, the inference speed in complex vegetation environments is stable at 28ms/frame, and the power consumption is controlled within 1.3W, providing a reliable technical tool for wild biodiversity research.

## 6. Conclusion

Through algorithm innovation and engineering optimization, this research effectively solves the problems faced by traditional identification methods such as poor real-time performance, high professional requirements, and heavy equipment. The system uses the improved yolov5 algorithm as the core detection framework, and through a series of technological innovations such as introducing an attention mechanism, optimizing the feature pyramid structure, and designing dynamic detection strategies, it has significantly improved the recognition performance of animals and plants in complex natural environments. The hardware adopts a modular design concept to achieve efficient collaboration of sensors, processors and display units, ensuring system performance while strictly controlling power consumption and volume.

Actual tests show that the system can effectively support multiple application scenarios such as field biological surveys, identification of agricultural pests and diseases, and nature education. Compared with existing methods, the system has improved the recognition accuracy of non-professional users and successfully helped scientific researchers discover multiple new species records. These results fully verify the practicability and reliability of the system, and demonstrate the huge potential of artificial intelligence technology in the field of ecological protection.

Future research will focus on improving the system's adaptability in extreme environments, exploring multi-modal perception and data fusion technologies, and improving the continuous learning mechanism to adapt to the discovery of new species and morphological changes. With the continuous optimization of technology and the expansion of application scenarios, this system is expected to become an important tool for biodiversity protection, laying the foundation for building an intelligent ecological monitoring network, and ultimately realizing the harmonious symbiosis of technology and nature.

## References

*[1] Xu Peng, Lu Zitong, Yang Yongjie. Improved YOLOv5 lightweight model suitable for embedded systems [J]. Electronic Design Engineering, 2025, 33(13):21-27+33.*

*[2] Zhang Fanbo, Ningbo, Kong Ningliang, et al. Research on surface defect detection technology based on machine vision [J]. New Technology and New Technology, 2025, (06): 74-80.*

*[3] Zhang Ming, Yan Fei, Wang Pengtao, et al. Automatic identification method of pavement diseases based on artificial intelligence [J]. Construction Technology (Chinese and English), 2025, 54(12): 129-133.*

*[4] He Zeshen. Research and experiment on pest monitoring system based on machine vision [D]. Heilongjiang Bayi Agricultural University, 2025.*

*[5] Li Ruohai. Case study of online detection algorithm based on artificial intelligence machine vision processing [J]. Automation Application, 2025, 66(10): 43-45+49.*

*[6] An Xiaodong, Li Yang, Liu Liu, et al. Review of fruit sorting algorithms based on machine vision [J/OL]. Automation Technology and Application, 1-12 [2025-10-13].*

*[7] Luo Sheng, Zhang Bin, Hu Yan, et al. Visual recognition method based on feature point matching [J]. Robotics Technology and Applications, 2023, (06): 24-28.*

*[8] Zhang Yuanmin, Li Dongdong, Huang Yong, et al. Research on the application of deep learning in the field of computer vision [J]. Technology and Innovation, 2024, (15): 40-42.*

*[9] Ke Ao, Wang Yucong, Hu Boyu, et al. Review of image-based wildlife detection and recognition [J]. Computer System Applications, 2024, 33(01): 22-36.*

*[10] Yang Tingting, Guo Zhixun, Lei Dingxiang, et al. Application of machine vision technology in agriculture [J]. Anhui Agricultural Science Bulletin, 2021, 27(18): 110-111.*