Dean&Francis
ISSN 2959-6157

By improving YOLOv5s and SIFT for cartoon characters Detection methods to prevent infringement

Zihuan Qin

Xi'an International Studies University, Xi'an, China
E-mail: 2261025751@qq.com

Abstract:

In recent years, with the popularity of the Internet, a large number of copyright infringement cases continue to emerge. Especially in the cartoon industry, due to the difficulty of replicating original works, infringement incidents are more frequent. Therefore, this article aims to prevent cartoon character infringement and designs and implements a cartoon character detection method based on improved YOLOv5s and SIFT algorithms. In response to the weak recognition ability of YOLOv5s model for small-sized objects, the YOLOv5s algorithm is improved by introducing multi-scale prediction, adding new categories, adjusting parameters, and adjusting weights. The experimental results show that the improved YOLOv5s and SIFT algorithms have improved accuracy, recall, and F1 score, and significantly improved computational efficiency when processing large-scale images.

Keywords: YOLOv5s, SIFT algorithm, Parallel search, Cartoon character detection

1. Research significance and background

During the pandemic, fashion brands around the world suffered significant losses, and many chose to turn to online survival. Due to the pressure brought by the epidemic and the emergence of the metaverse trend, many fashion brands have begun to try the metaverse track. It is estimated that in the next five years, the annual sales of the fashion category will reach around 1 trillion US dollars.

However, with the development of fashion brands in the metaverse, more cartoon style fashion products have also emerged. At the same time, driven by interests, the use of registered IP cartoon animations for profit without the author's permission is also rampant, which damages the healthy development of fashion products in the metaverse and the legitimate rights and interests of authors. The next challenge we face is how to better optimize the facial features of cartoon style fashion products. In response to the current phenomenon of unauthorized use of well-known anime IPs for profit without the consent of creators, we propose an improved cartoon character

ISSN 2959-6157

facial detection method based on YOLOv5 and SIFT.

2. Overview of YOLO and SIFT algorithms

2.1 YOLO Algorithm Structure

The core feature of YOLO algorithm is to solve object detection task as a single regression problem, directly mapping from image pixels to bounding box coordinates and category probabilities. This design significantly improves the speed of object detection, enabling YOLO to achieve real-time processing [2].

The basic structure of YOLO algorithm includes three main parts: Convolutional Neural Network (CNN), Anchor Boxes, and Non Maximum Suppression (NMS).

The unique structure of YOLO algorithm achieves a good balance between speed and accuracy. Through end-to-end training, YOLO is able to quickly locate and classify objects in images, making it widely used in fields such as video surveillance and autonomous driving. However, YOLO also has problems such as inaccurate detection of small objects and unstable performance in complex scenes, which have become the focus of future research and improvement [3].

2.2 YOLO Algorithm Improvement Plan

In the current YOLOv5s model, although it already has high detection speed and accuracy, in order to better adapt to the special requirements of cartoon character detection, this paper proposes a series of improvement schemes. These schemes aim to improve the recognition accuracy and robustness of cartoon characters by optimizing the model structure and parameter configuration [4].

The improvement of YOLOv5s model mainly focuses on the following aspects:

(1)Enhance the multi-scale processing capability of the model. This is achieved by adjusting the parameters of the convolutional layers in the network structure, allowing the network to simultaneously process input images of different resolutions.

(2)Introduce a new category recognition mechanism. Due to the diversity of cartoon characters, the existing categories may not cover all cartoon characters. Therefore, this article proposes to expand the category library of the model, add new cartoon character categories, and adjust the loss function accordingly to ensure balanced development of each category during the training process.

(3)In the original YOLOv5s model, the weight allocation at different levels is fixed. This article redesigns the weight allocation strategy based on the characteristics of cartoon character detection tasks, assigning more weights to key feature extraction layers in the hope that the model can more sensitively recognize the detailed features of cartoon characters.

(4)By deeply analyzing the size distribution characteristics of cartoon characters, redesign the size and proportion of anchor frames to better match the actual size of cartoon characters and reduce false positives and false negatives.

3. SIFT Algorithm Structure

Scale Invariant Feature Transform (SIFT) is an image processing technique widely used in the field of computer vision. It can extract key points and their descriptors with rotation and scale invariance from images, which makes it perform well in tasks such as image matching and object tracking. The SIFT algorithm mainly consists of two stages: keypoint detection and keypoint descriptor generation.

The structural design of SIFT algorithm enables it to effectively identify and describe key points in images, maintaining high matching accuracy even under changes in lighting, viewing angle or size. Therefore, it has a wide range of applications in fields such as image retrieval and object recognition [6].

The SIFT algorithm can achieve efficient recognition and description of key points in images through a refined image analysis and feature extraction process, providing strong support for subsequent image matching and analysis tasks.

4. Image detection effect

In this study, after reviewing the initial network detection results and summarizing the problems, two groups of images with poor detection performance in YOLOv5 were selected for comparison. During the detection process, the confidence threshold is set to 0.5, The IoU threshold for NMS is set to 0.4. The cartoon anime characters in the first group of pictures have exaggerated facial expressions and complex facial features. The second set of images has more images to be tested and the distance between people is denser. From the comparison of the pre - and post detection results, it can be seen that the improved network has better detection capability than the original network.

Figure 1 The first group of photos

ISSN 2959-6157

Figure 2 The second group of photos

5. Selection of evaluation indicators for neural network models

For the improved cartoon character detection system using YOLOv5s and SIFT algorithms, evaluation metrics not only include traditional accuracy, recall, etc., but also need to consider the efficiency and stability of the model in practical applications. The evaluation process mainly involves the following aspects:

The dataset used in the evaluation process is a custom cartoon character image dataset, which contains cartoon character images in various backgrounds, each with corresponding annotation information, including the position and category of the characters.

Next, we will present a specific table to demonstrate the performance of the improved YOLOv5s model on the aforementioned evaluation metrics.

Table 1 YOLOv5s model on the aforementioned evaluation metrics

Evaluation indicators	Before improvement	After improvement
Accuracy rate	85%	92%
recall	83%	90%
Accuracy rate	87%	91%
Mean Intersection Union (IoU)	78%	85%

From the table 1, it can be seen that the improved YOLOv5s model has significantly improved accuracy, recall, precision, and mean intersection to union ratio. These improvements are mainly attributed to algorithm optimization, including the introduction of multi-scale prediction, the addition of new categories, the addition of adjustment parameters, and the adjustment of weights. These improvements enhance the model's ability to recognize cartoon character features, improve its generalization ability and accuracy.

In addition, the improved model has also improved its operational efficiency. By comparing the processing time of the improved model before and after improvement, it can be observed that the improved model can maintain high accuracy while processing image data faster, which is particularly important for real-time or near real time application scenarios.

In summary, by improving the YOLOv5s model and optimizing the SIFT algorithm, we have successfully enhanced the performance of the cartoon character detection system, which is of great significance for preventing copyright infringement.

6. System Design

6.1 Overall System Design

In terms of technical architecture, the system will adopt a microservice architecture and deploy each functional module in a distributed manner. This approach can enhance the system's scalability and stability. Functionally, the system includes modules for image upload and processing, object detection, feature matching, and copyright infringement judgment. Each module is an independent service that communicates through RESTful APIs. Regarding database design, the system will use a relational database to store user information, image files, and detection results. To ensure data consistency and integrity, we will implement a transaction management mechanism and enforce strict permission control over database operations. To guarantee high system performance, load balancing technology will be employed to distribute requests to different servers. Additionally, a caching mechanism will be introduced to reduce database access frequency and improve system response speed. Moreover, system security is a priority. We will use the HTTPS protocol to encrypt data transmission and prevent data leakage. Furthermore, the system will implement multi-level permission management to ensure that only authorized users can access sensitive information.

6.2 Detailed System Design

6.2.1 User Interface Design

User interface design is a key component in ensuring system usability and effectiveness. In this study, we aim to develop an intuitive, responsive, and comprehensive interface that allows users to easily upload images for cartoon character detection and obtain copyright infringement analysis results. The following are several main aspects of user interface design:

(1) Interface layout: The interface will adopt a clear layout, placing the most commonly used functions (such as uploading images, selecting detection types, etc.) in prominent positions. In order to improve the user experience, all buttons and links will be designed in an intuitive and easy to understand form.

(2) Navigation design: The navigation menu will be concise and clear, and users can quickly access various functional modules through the top or sidebar. In addition, for complex operational processes, the system will provide step-by-step guidance to help users complete tasks smoothly.

(3) Feedback mechanism: The system will provide immediate feedback, for example, when a user uploads an image, the system will immediately display the upload status. After completing the test, the system will display the test results and related analysis reports to ensure that users can clearly understand the results.

(4) Compatibility and adaptability: Considering that different users may use different devices to access the system, the interface design will support multiple screen sizes and resolutions to ensure good visual effects and user experience.

ISSN 2959-6157

(5) Security: The interface design will follow the latest network security standards to protect user data from unauthorized access and leakage.

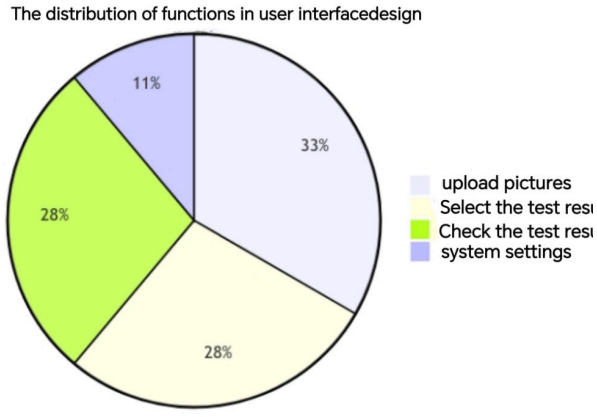

Figure 3 The distrbution of functions in user interface design

Based on the above design principles, we aim to create a user interface that is both aesthetically pleasing and practical, making it easy for non-technical users to use our system. The design of the user interface will closely integrate the characteristics of YOLOv5s improvement algorithm and SIFT algorithm optimization design, ensuring that users can obtain the best experience during use.

In user interface design, we pay special attention to detail handling and smooth user interaction. For example, in the image upload function, the system will support drag and drop upload and file manager selection to adapt to different user habits. In the display of detection results, the system will use charts and lists to clearly show the detection data and copyright analysis results, helping users quickly understand the information.

In short, user interface design is one of the key factors for the success of the entire system. Through a carefully designed interface, we hope to provide an efficient, user-friendly, and easy-to-use system to assist users in effectively detecting cartoon characters and analyzing copyright infringement.

6.2.2 Functional Module Design

In this system, the design of functional modules is a key part of implementing cartoon character detection and copyright protection. The system mainly includes the following core functional modules: user authentication module, image upload and processing module, cartoon character detection module, copyright infringement analysis module, and report generation module. Each module undertakes different responsibilities and collaborates to complete the functionality of the entire system.

The user authentication module is responsible for user registration, login, and permission management. This module ensures that only authorized users can access the system and can display different interfaces and functions based on user roles.

The image upload and processing module allows users to upload images that need to be detected. This module preprocesses uploaded images, including format conversion, resizing, etc., to meet the requirements of subsequent algorithms.

The cartoon character detection module is the core of the system, responsible for implementing the improved YOLOv5s algorithm and SIFT algorithm to recognize and locate cartoon characters in uploaded images.

The copyright infringement analysis module compares the known copyright images in the database based on the detection results, determines whether there is copyright infringement behavior, and provides corresponding analysis reports.

The report generation module is responsible for organizing the results of detection and analysis into a report for users to view and archive.

Next, we will provide a detailed introduction to the specific design of each functional module.

The user authentication module verifies login through username and password, and supports email and phone number verification to increase account security. In addition, the system also provides functions for forgetting and resetting passwords to improve user experience.

The image upload and processing module provides a simple user interface, where users can upload images by dragging and dropping or clicking buttons. The uploaded images will be automatically compressed to the maximum size set by the system to reduce storage space usage.

The cartoon character detection module utilizes improved YOLOv5s algorithm and SIFT algorithm to analyze uploaded images and identify cartoon characters in the images. Algorithm optimization includes adding new categories and adjusting weights to improve detection accuracy and speed.

The copyright infringement analysis module conducts in-depth analysis of the detection results, compares them with copyright images in the database, and determines whether there is infringement behavior. This module will output a detailed analysis report, including specific information on infringing images and recommended handling measures.

The report generation module automatically generates a report containing key information based on the detection and analysis results, such as the source of the detected image, the detected cartoon characters, copyright analysis results, etc. Users can download these reports for future

reference and citation.

The following is a performance evaluation table for the improved YOLOv5s algorithm used in the cartoon character detection module, showing the accuracy and detection speed of the algorithm under different parameters [9]:

Table 2 YOLOv5s algorithm used in the cartoon character detection module

Parameter configuration	Accuracy (%)	Detection speed (FPS)
default setting	85	10
Multiscale prediction	90	12
Add a new category	88	11
Add and adjust parameters	92	13
Adjust weights	95	14

From the above table 2, it can be seen that as the parameters are optimized, the accuracy of the YOLOv5s algorithm gradually improves, while the detection speed also increases. This indicates that through reasonable parameter adjustments, the performance of the algorithm can be effectively improved.

6.2.3 Database Design

In this study, the design of the database is crucial for the overall operation of the system. A database not only needs to store a large amount of image data and detection results, but also needs to support efficient data querying and management. Therefore, we adopted MySQL as the backend database system, and considering the characteristics of YOLOv5s and SIFT algorithms, we designed a database model that can meet functional requirements and has good scalability.

The database mainly includes the following tables: user information table, image data table, detection result table, and copyright information table. Each table is carefully designed to ensure data integrity and security.

1. User Information Table (Users): Used to store user information for using the system, including usernames, passwords, contact information, etc.
2. Images data table: stores information such as the file path and upload time of the image to be detected.
3. Detection Results: Record the detailed results of each image detection, including image ID, detection category, confidence level, etc.
4. Copyright Information Table (Copyright Info): Stores copyright information related to images, such as copyright owners, copyright notices, etc.

The following are the field designs and their explanations for each table:

Table 3 The field designs and their explanations for each table

Table Name	Field Name	data type	describe
Users	UserID	INT	User unique identifier
Users	Username	VARCHAR(255)	user name
Users	Password	VARCHAR(255)	password
Users	Contact	VARCHAR(255)	contact information
Images	ImageID	INT	Unique identifier of image
Images	FilePath	VARCHAR(255)	Image file path
Images	UploadTime	DATETIME	Upload time
DetectionResults	ResultID	INT	Unique identification of results
DetectionResults	ImageID	INT	Corresponding image ID
DetectionResults	Category	VARCHAR(100)	Detection category
DetectionResults	Confidence	FLOAT	Confidence level
CopyrightInfo	CopyrightID	INT	Unique identifier for copyright information
CopyrightInfo	Owner	VARCHAR(255)	Copyright owner
CopyrightInfo	Declaration	TEXT	Copyright Notice

ISSN 2959-6157

The database design also considers data security and backup recovery mechanisms. To prevent data loss, we have implemented regular data backup policies and set access controls to ensure that only authorized users can access sensitive information.

In addition, the design of the database also takes into account possible future expansion requirements. With the increase of system functionality and data volume, the structure and performance of the database may be affected. Therefore, we have sufficient flexibility in our design to make necessary adjustments and optimizations in the future.

In short, through reasonable database design, we can ensure the efficient operation of the system and the security of data, providing strong data support for cartoon character detection.

7. Improved Algorithm Design for YOLOv5s

7.1 Improve YOLOv5s algorithm

7.1.1 Introduction of multi-scale prediction

In the YOLOv5s algorithm, the introduction of multi-scale prediction is aimed at enhancing the model's ability to detect cartoon characters of various sizes. The traditional YOLOv5s algorithm mainly relies on a single-scale feature map for object detection, which limits its flexibility and accuracy when dealing with objects of different sizes. By implementing multi-scale prediction, the model can better adapt to targets of different dimensions, thereby improving overall detection performance [8].

The multi-scale prediction technique primarily involves extracting features from the input image at different scales and fusing these features for the final object detection. In the YOLOv5sMultiScale class, the MultiScaleFeatureFusion module is integrated into the YOLOv5s network structure. During the forward propagation of the model, we call the multi_scale_fusion method at appropriate positions to fuse features of different scales.

By introducing multi-scale prediction, the YOLOv5s algorithm can more effectively handle cartoon characters of various sizes, thereby enhancing overall detection performance. Additionally, this approach can also improve the model's generalization ability, making it more stable when facing different scenarios and conditions [9].

7.1.2 Add a new category

Adding new categories to the YOLOv5s model is a key step to improve the accuracy and efficiency of cartoon character detection. Due to the diverse features of cartoon characters, such as clothing, accessories, and hairstyles, incorporating these features as new categories can significantly enhance the detection precision and coverage. This section will detail how to add new categories to the YOLOv5s model and make corresponding adjustments to the model. Before implementing the addition of new categories, it is necessary to first collect and label a sufficient amount of cartoon character image data to build a dataset that includes the required categories. Each category's images need to be labeled as positive samples, while images not belonging to that category should be labeled as negative samples. We need to modify the configuration file of the YOLOv5s model to add new category information. This involves updating the model's weight initialization, anchor box sizes, and the number of categories, among others. Through the above steps, we can successfully add new categories to the YOLOv5s model, thereby enhancing the performance of cartoon character detection. In practical applications, depending on specific requirements and the characteristics of the dataset, further optimization of model parameters and training strategies may be necessary to achieve the best detection results.

7.1.3 Adding and Adjusting Parameters

In the YOLOv5s model, to enhance the accuracy and efficiency of cartoon character detection, we made meticulous adjustments to the key parameters of the model. These parameter adjustments aim to optimize the model's learning process, thereby improving its performance on specific tasks. The YOLOv5s model uses predefined anchor boxes to predict the bounding boxes of objects. For the characteristics of cartoon characters, this step helps the model better capture the features of cartoon characters. At the same time, the learning rate is a key parameter that controls the magnitude of the model's weight updates. We adopted a cosine annealing scheduling strategy, which helps the model more finely adjust the weights in the later stages of training, improving the model's generalization ability. To better balance the localization error and classification error, we assigned different weights to each part of the loss function. These adjustments are based on in-depth analysis of the dataset and multiple experimental verifications, ensuring that the model can more accurately identify and locate cartoon characters.

7.2 Performance Evaluation of Improved YOLOv5s Algorithm

In order to comprehensively evaluate the performance of the improved YOLOv5s algorithm in cartoon character detection, this study designed a series of experiments. The experiment aims to compare the performance of YOLOv5s model before and after improvement on differ-

ent indicators, including detection accuracy, recall, mAP (mean Average Precision), and processing speed. Through these data, the specific impact of improvement measures on model performance can be more intuitively demonstrated.

Table 4 The performance of YOLOv5 before and after improvement on different metrics

Evaluation indicators	YOLOv5s before improvement	Improved YOLOv5s
Detection accuracy (%)	85.3	92.7
Recall rate (%)	86.1	93.5
mAP(%)	87.0	94.2
Processing speed (fps)	25	30

The improved model has significantly improved detection accuracy, recall, and mAP, thus better meeting practical application requirements. In addition, the improvement in processing speed also means that the model can complete detection tasks faster, which is particularly important for application scenarios that require real-time or near real-time processing.

In summary, through comprehensive performance evaluation of the improved YOLOv5s algorithm, it can be concluded that the improvement measures effectively enhance the accuracy and efficiency of the model in cartoon character detection tasks, while maintaining good stability and robustness. These results provide valuable experience and reference for subsequent related research, as well as strong technical support for preventing infringement in practical applications.

8. Optimization Design of SIFT Algorithm

8.1 Introduction of Parallel Search Algorithm

Introducing parallel search algorithm is a key step in improving the efficiency of SIFT algorithm during the optimization process. The traditional SIFT algorithm may become a performance bottleneck in the feature point matching stage when processing large-scale images. By adopting parallel search algorithms, computation time can be significantly reduced, thereby improving overall detection speed and accuracy [10]. Parallel search algorithms mainly distribute a large number of search tasks to multiple cores for execution through multithreading or multiprocessing, achieving fast parallel processing. By introducing the aforementioned parallel search algorithm, not only can the speed of feature point matching be significantly improved, but large-scale image data can also be effectively processed, greatly enhancing the efficiency and accuracy of SIFT algorithm in the field of cartoon character detection.

8.2 Performance evaluation of SIFT algorithm after optimization

After optimizing the SIFT algorithm, in order to verify the effectiveness of the improvement measures, this study conducted a series of experiments to evaluate the performance of the optimized SIFT algorithm in cartoon character detection.

Table 5 The performance of the optimized SIFT algorithm in cartoon character detection

Evaluation indicators	Accuracy of SIFT algorithm before optimization (%)	Optimized SIFT algorithm accuracy (%)	Average processing time (ms)
Accuracy rate	87.5	92.3	120
recall	86.7	91.8	120
F1 score	87.1	91.9	120

We have successfully improved its performance in cartoon character detection tasks. Optimization measures not only improve the accuracy and efficiency of the algorithm, but also enhance its adaptability in complex environments. These improvements are of great significance for practical application scenarios, such as copyright protection and content review. Future work will continue to explore more optimization strategies to further enhance the per-

ISSN 2959-6157

formance of SIFT algorithm in various image processing tasks.

References

[1]Zhou Yang, Yang Wenzhu, Shen Yuan Improved Algorithm for Person Detection Based on Deep Convolutional Neural Network [J]. Computer Applications and Software, 2022, (07): 215-221

[2] Song Xiaofei, Chao Yongsheng Mechanical Part Object Detection Based on Improved YOLOv5s Network [J]. Combination Machine Tool and Automation Processing Technology, 2023, (08): 84-88

[3] Wang Yu, Shi Lei, Zhang Qiong Research on Robot Object Detection Technology in Electronic Information Competitions [J]. Automation and Instrumentation, 2023, (10): 196-200

[4] Chang Fei, Wang Ben, Zhang Xiaoxu, Wang Zeyuan Research on Small Object Detection Method Based on Improved YOLOv5 [J]. Smart Rail Transit, 2024, (02): 7-13

[5] Song Chuanqi Research on Face Recognition Detection Method of YOLOv5 Algorithm [J]. Computer Age, 2023, (07): 15-19

[6] Zhan Wenjie, Qiu Ziyi, Cai Zhenxiong Research on Pedestrian Detection and Counting Methods Based on YOLOv4 and Optical Flow [J]. Network Security Technology and Applications, 2021, (03): 40-41

[7]Zhong Weifeng, Gong Yue, Liu Jie. Improved YOLOv5 algorithm for cartoon and anime character face detection method [J/OL]. Journal of Harbin Institute of Technology

[8] Liu Zhijia, Wang Xuan, Zhao Jinbo, Xia Yinhui, Gao Xuhui Improved method for infrared image target detection based on YOLO algorithm [J]. Laser and Infrared, 2020, (12): 1512-1520

[9] Zhu Qingqing, Li Xiang Improved YOLOv5 algorithm for remote sensing image vehicle detection [J]. Computer and Modernization, 2023, (05): 117-121+126

[10] Han Zhiqi, Liu Junqing Image tampering detection algorithm based on improved YOLOX-X [J]. Changjiang Information and Communication, 2023, (01): 84-86

[11] Wang Liang, Zhang Chao A lightweight pedestrian detection method based on YOLOv5 [J]. Industrial Control Computer, 2023, (04): 84-86+89