Family service robot intellisense and autonomous navigation technology research

Pengfei Guo[a],Ziheng Hou[b],Wenhai Liu[c],Mengshuai Cao[d],Wenchao Wu[d]
North China Electric Power University (Baoding), Baoding 071066,Hebei, China

[a]220191110505@ncepu.edu.cn,[b]1833132519@qq.com,[c]2207237511@qq.com,[d]3161597329@qq.com,[e]2033541424@qq.com

Abstract—The research on intelligent sensing and autonomous navigation technology of home service robot is based on the research of IFlytek home service robot. The robot can conduct positioning and path planning through odometry, liDAR, imu and other related technologies, and complete autonomous navigation and simple obstacle avoidance operation. Based on IFlytek's voice technology, the robot can realize the recognition of specific words, and then complete the task of speech recognition and start. Based on the computer vision algorithm to complete the object detection task to determine the type of room, so as to better carry out the corresponding service, to achieve the purpose of room exploration.

Key words—Natural Language processing; computer vision; SLAM; path planning; Intelligent perception

Ⅰ. INTRODUCTION

In order to perform the family service task better, the family service robot needs to carry on the advance cognitive work to the corresponding family environment. For example, for different families, their layout is not the same, so they need the judgement for different home environment and objects. [1] . We start by assuming that different life scenarios can be in three rooms. There are eight kinds of printed image board, the robot from the returned images, to identify the corresponding categories.Based on this, the types of the three rooms can be determined. Therefore, the correct recognition of the image board is the premise of judging the type of room. In practical applications, intelligent robots can determine the type of a room based on their own cameras, and thus perform different tasks in different rooms. In this paper, autonomous navigation, voice control, robot and intelligent home connectivity and other issues are studied, and through the practice of the feasibility of the study [2] .

There are three main problems to be solved in this research: ①how to locate, navigate and plan the path in the environment with obstacles; ② how to recognize specific objects in the environment; ③ how to carry out speech detection and synthesis.

The research's innovation is that the robot can fully integrate natural language processing, computer vision, SLAM, path planning and other technologies,as well as the combination of these technologies.

Ⅱ. OVERALL SCHEME DESIGN

2.1 the design idea of the identification section

2.1.1. Collection of data sets

(1) The team first zoomed, rotated, and removed the white background from the 32 sample images, then pasted them onto the background images of the other images. Finally, the sample image is processed by Gaussian blur and changing brightness, contrast and saturation to get the data set.

(2) Image tagging the robot can get the image automatically without further tagging by manual.

2.1.2 Identification model

The recognition model used by the team is the Yolov5 algorithm, which is characterized by：① the use of code generation data sets, saving the cost of human tagging; ② when the actual effect of model recognition is not good, you can automatically modify the code, automatic generation of new data sets, saving most of the time and effort; ③use the model yolov5n, CPU speed

faster, predicted 15 photos to spend less than 6s. In the actual prediction, the model can correctly judge the vast majority of images, and misclassified the environment in very few extreme conditions.

Ⅲ. ALGORITHM FUNCTION

3.1 The navigation section

In the process of navigation, the first step is to use Gmapping [3] algorithm to establish the environment map at the same time of localization, and to complete the map establishment by moving constantly, which is the premise of robot navigation and path planning.

The robot completes the path planning through TEB + Dijkstra algorithm, planing out the global path, and completes the local path planning on the basis of the global path, making the robot motion conforms to the dynamic constraints and can avoid the obstacles that do not appear on the static map. Among them, the parameters of Teb are the most important, which determine the performance of TEB and the effect of robot motion. After fully understanding the principle of TEB algorithm, we examine its motion through survey parameters.

3.2 The identification part

Data set Generation Algorithm: the data of our data set is obtained by preprocessing 32 target images by code and pasting them onto other images (background) . The target image is the pre-processed image of 32 pdfs. Background image is the original image pasted by the target image, and its content is independent of all the target images. Non-uniform scaling means that the aspect ratio

of the image changes when the image is modified. Table 1and Table 2 generate a list of parameters for the data set, as shown in the following figure.

Table 1 The Data Set Generates A List Of Parameters

parameter name	parameter s mean	default
one2many_num	An image generates a multiples of the image	100
BG_im_nums	The maximum number of targets contained in an image	3
max_try	The maximum number of attempts to paste the target image	5
max_IMU	The upper limit of the IMU of the target image with all other pasted target images	0.3
change_range	Two-dimensional list, control the scaling and rotation of posted images, [[minimum, maximum], [minimum rotation Angle, maximum rotation Angle]]	[[0.15,0.9],[-30,30]]

Table 2 The Data Set Generates A List Of Parameters

·yolov5

Figure 1. Backbone procedure chart

parameter name	parameters mean	default
locate_type	Data set label format	yolo
h_size	The height of the final image	640
remove_background	The probability of removing the pasted image background	0.6
enhance	bool Indicates whether to change the contrast, brightness, saturation, and Gaussian blur of the image	True
max_blue	Gaussian blur parameter, PIL. ImageFilter. GaussianBlur (radius) of the radius of maximum	5
min_enhance	Minimum value of the PIL image enhancement parameter	0.5

The input includes three parts: Mosaic data enhancement, image size processing and adaptive anchor frame calculation. Mosaic data enhancement can combine four images to enrich the image background, and the image size processing adaptively adds the least black edge to the original image with different lengths and widths, and uniformly scales to the standard size [4] adaptive anchor frame calculation on the basis of the initial anchor frame, the output prediction frame is compared with the real frame, and the difference is calculated and then updated in the reverse direction, and the most suitable anchor frame value is obtained by iterating the parameters continuously.

Backbone consists mainly of the Bottleneck CSP and Focus modules. The Bottleneck CSP module improves the convolutional neural network learning performance while significantly reducing the amount of computation, [5] and the Focus module slices the image to quadruple the input channel.The lower sampling feature map is obtained by one convolution, which reduces the computation and speeds up the process of sampling. [6] Neck adopts the structure of combining FPN and PAN, which combines the conventional FPN layer with the bottom-up feature pyramid, and can realize the fusion of the extracted semantic features and location features.At the same time, the main layer and the detection layer are fused to make the model acquire more abundant feature information. Head outputs a vector that has the category probability of the target object, the object score, and the location of the object's bounding box. The detection network consists of three detection layers, and the feature maps of different sizes are used to detect objects of different sizes.After each detection layer outputs the corresponding vector, the prediction boundary box and category of the target in the original image are generated and marked. [7]

The Yolov5 configuration parameter

The team used the identification model yolov5n.For the parameters of the model itself, the team only modified the nc (LRB-number of categories) to 8 and the rest of the parameters are official configuration.

The revised parameters of the training are as follows:

1. EPOCH = 100

2.Optimizer: Adam

3. Imgsz : 320

3.3Room recognition algorithm

The room recognition algorithm takes into account the failure of the recognition. If the recognition error occurs, certain preventive measures are set, but the recognition failure may also occur. When the recognition fails, the unrecognized room type can be randomly returned based on the recognized room type. As shown in Table 3.

Table 3

A list of room items

dining room	table chairs	tableware	food	people
Bedroom	people	pet	bed	table chair

Living room	sofa	pet	table chair	TV

The design algorithm is as follows:

3.3.1Single room recognition algorithm：

```
quick_set = {"tableware","food","bed","sofa","TV"}
quick_out = {
    "tableware": "Dining room",
    "food": "Dining room",
    "bed": "Bedroom",
    "sofa":"Living room",
    "TV":"Living room"
}
slow_out = [[
    {"people","pet","table chair"},"Bedroom"
]]
```

Figure 2. Diagram of Constant declaration

Meaning Description:quick set: If the item belongs to the set, it can be directly returned to the room type according to quick out; slow out: Category mapping for longer combinations.

Parameter Description:

1. class list: The type is a list and can be empty.
2. id class name: python dictionary object {int: class name}, used to map numeric type elements in class list to character type category names.

The algorithm flow is as follows:

1. Determine the type of each element in the class list. If it is float, convert it to int, and then convert it to str (class name). If int, convert to str (category name)
2. Convert class list to collection class set.
3. Iterate over each element in the class set to see if it belongs to the quick set, and if it does, return the room type directly based on the quick out map.
4. If 3 is not returned, check whether class set is equal to slow out[i,0] according to traversal slow out. If so, return the corresponding category slow out[i,1]

If 4 is not returned, the recognition failed and None is returned.

3.3.2 Overall room recognition algorithm:

Parameter Description:

1. ans: For elements in ans, the format is [image path,[[probablit 0, class 0], [probablit 1, class 1],[probablit 2, class 2],...]
2. image path:
3. probablit i, class i: The probability of the model predicting class i in this picture is probablit i
4. save filename: txt file path to save the result,
5. id class dict: {int: class name} dictionar
6. class room dict: {class:[room 0, room 1]} dictionary
7. room name list: Room name list
8. threshold: Probability threshold

3.3.3 Algorithm flow:

1. According to the image path in each element of ans, determine the room serial number id represented by this element. According to the id class name dict, convert each class i in ans to the corresponding category name and save it in the category list of the corresponding serial number room.
2. Implement algorithm for each room separately: single room recognition algorithm
3. If there is an unrecognized room, the result will be randomly changed to an unrecognized room, so that the room type in the final result is 3.

Ⅳ. THE IMPLEMENTATION PROCESS

Part of speech

1. Familiar with Linux SDK

If the device is started for the first time and the udev rule for the microphone has not been configured, run the following command to configure the device. If the Udev rule has been configured, skip this step. Note: xf mic.rules is located in the root directory. Then run the following command to restart the udev service:

$ sudo service udev restart

Install the sound card library and audio player library

Dynamic library configuration

Generate executable files (basic function demonstration)

2. ROS Feature pack

Configuring the offline engine

Download iFlytek sdkhttps://www.xfyun.cn/ Create a new app --> Console --> Download Offline Speech Recognition SDK(select Linux MSC)

Ensemble Methods

Replace the ros interface package:

Replace the user's own common.jet file with vvui ros-master/xf mic asr offline/config/ms/res /asr. Replace appid with common.jet in vvui ros-master/xf "appid" in the mic offline params.yaml file in mic asr offline/config, then compile again.

Part of identification

DataSet

As mentioned above, the data set used by this team is code generated, and the original image designed is 32 sample images provided by the official and the background image crawled from the search engine. The data set was generated for three times with 3200 pieces each time, and the parameters and algorithms of each time were improved to some extent based on the last time.

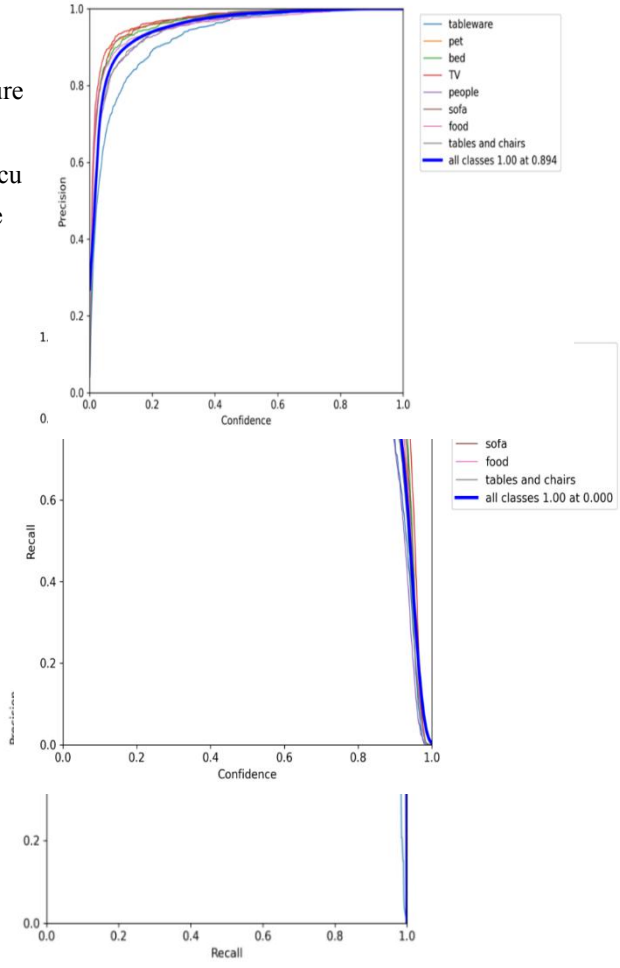
After the data set was generated, the model was run on the data set. The data set was divided into the training set and the verification set three times according to 8:2. The next step is the actual effect of the model, and then the generation algorithm and parameters of the data set are adjusted again according to the actual effect to generate a new data set. The model is then used to run both old and new data sets together or select parts of them.

Initial data set generation algorithm 1.0:

1. The original generation algorithm has
2. background map was added: "sky", etc., and the shrinkage multiple and rotation Angle were changed.
3. Perspective transformation in image processing is considered, but used on reshaping simulation, which is to transform the aspect ratio of the sample image.

. DATA ANALYSIS

identification division



igure 3. P_curve

considered cutting the sample image, removing the white background of the sample image, and then shrinking and pasting it onto the downloaded background image.

2. At this time, the scaling is consistent, and the aspect ratio of the picture does not change before and after.

Data generation algorithm 1.1:

1. The second data set generation algorithm considers the brightness, contrast, saturation and clarity of the image. PIL database was used for corresponding transformation of the data set, and the clarity was modified using Gaussian fuzzy algorithm to increase the robustness of the trained model.

2. At the same time, a new parameter "remove background" is added, indicating that for the sample image img i, the probability "remove background" will be used to remove its background

data set generation algorithm 1.2:

1. In the third data set generation process, a new

igure 4. PR_curve

Figure 5. R_curve

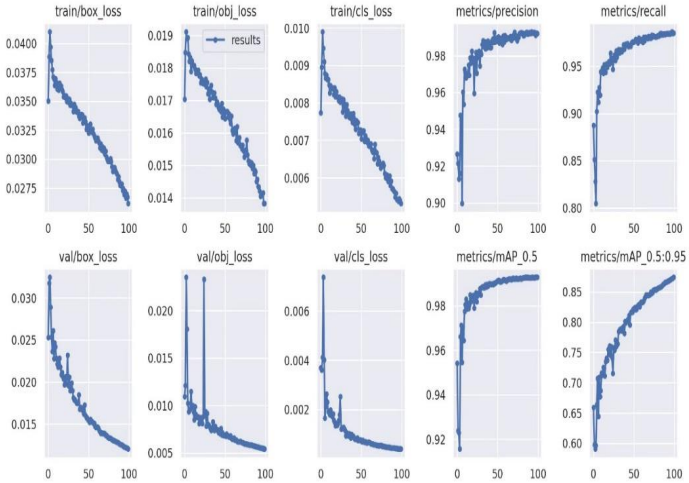


Figure 6. Data transformation diagram

Comparison of BG-pred during training

Figure 7. BG sketch map

Figure 8. pred effect picture

It can be seen from the verification set and the actual results that the model has a strong discrimination ability.

Ⅵ. CONCLUSIONS

The recognition model of the family service robot performs well in the actual test and can correctly identify most of the targets. In the decision-making, we try our best to give full play to the car's performance, improve efficiency, save time, and achieve a balance between speed and recognition.

REFERENCES

[1] Yiu Yat-fai, Chan Man-pak, Chan Kai-lai, ng Pui-leung. Construction and application of knowledge map
of Home Service Robot [J/OL] . Acta Beijing University of Posts and Telecommunications: 1-7[2022-10-09] .

[2] Jiang Hengxin, Yan Xiaole, Wang Lu. Research on control design of Voice Home Service Robot based on ROS system [J] . Innovation in Science and Technology, 2022(04) : 158-161.

[3]Grisetti G, Stachniss C, Burgard W.Improved techniques for grid mappingwith raoblackwellized particle filters[J].IEEE Transactions on Robotics, 2007,23(1):34-46.

[4] He Ying, Chen Ding Hao, Peng Lin. Study on the detection algorithm of economic forest pests based on improved Yolov5 Model [J] . Chinese Journal of agricultural machinery chemistry, 2022,43(04) : 106 -115.

[5] Liu Ying, Zhang Xuesong, Liu Jiyue. Design of mask-wearing detection system based on Yolox target
detection algorithm [J] . Microcontroller and embedded system applications, 2022,22(06) : 87-91.

[6] Zhai Zhiyu. Study on recognition of pedestrian abnormal behavior in parking lot based on multi-feature fusion [D] . Teacher: Hou Beiping. Zhejiang University of Science and Technology, 2022.

[7] Yang Qisheng, Li Wenkuan, Yang Xiaofeng, Yue Linxi, Li Haifang. Improved Yolov5 method for detecting the growth state of Apple and peanut [J] . Computer Engineering and applications, 2022,58(04) : 237-246.