UChile Homebreakers 2017 Team Description Paper

Luz Martínez, Rodrigo Muñoz, Gonzalo Olave, Giovanni Pais, Gustavo Hernan, David Gomez, Leonardo Garrido, Diego Campanini, Pedro Orellana, Patricio Loncomilla, and Javier Ruiz-del-Solar

Department of Electrical Engineering - Advanced Mining Technology Center
Universidad de Chile
http://robotica-uchile.amtc.cl/bender-index.html

Abstract. The UChile HomeBreakers team is an effort of the Department of Electrical Engineering and the Advanced Mining Technology Center of the Universidad de Chile. The team has participated in the RoboCup @Home league since 2007, and its social robot, Bender, obtained the @Home Innovation Award in 2007 and 2008. As a team with strong expertise in robot vision, object recognition, and human-robot interaction, we believe that we can provide significant inputs to the league. This year our main research focus is the incorporation and development of Deep Learning in Computational Vision for domestic robots. Deep Learning has been achieved by the use of region proposals and object classification with CNN algorithms. Furthermore, a semantic segmentation system based on pre-trained CNN features has been developed for place recognition. In addition, indirect object search capabilities were added using a bayesian based methodology. Another important development in Bender is a completely new head that allows it to more easily display emotions. The software includes new developments of face detection/recognition, and the recognition of facial features like age and gender, all using deep learning. Finally, improvements in speech recognition have been made using PocketSphinx.

1 Introduction

The UChile HomeBreakers team is focused on fostering the research of mobile robotics and other robotics-related areas of investigation (e.g. human-computer interactions, machine learning, amongst others). The main motivation of the team is to work on the continuous development of technologies for service robots -whether it is software or hardware- to help people. One of the best ways to accomplish these goals is participating in international robotics competitions, in which the team can acquire and share knowledge with other research groups, and test the quality of the developed technologies. In addition to testing and showing progress achieved in competitions, the team conducts industrial projects, publishes papers, and provides educational activities for high school and university students.

2 Background

The UChile HomeBreakers team has participated in RoboCup competitions since 2003 in different leagues: Four-legged 2003-2007, @Home in 2007-2015, Humanoid in 2007-2010, and Standard Platform League (SPL) in 2008-2016. Universidad de Chile team members have served the RoboCup organization in many ways (e.g. TC member of the @Home league, Exec Member of the @Home league, and co-chair of the RoboCup 2010 Symposium). One of the team members is also an organizers of two special issues on ”Domestic Service Robots” of the Journal of Intelligent and Robotic Systems. As a RoboCup research group, the team believes that its contribution to the RoboCup community is not restricted to the participation in the RoboCup competitions, but that it should also contribute new ideas. In this context, the team has published a total of 36 papers in the RoboCup Symposium (see Table 1); in addition to many other publications about RoboCup related activities in international journals and conferences. Amongst the most important scientific achievements of the group are obtaining three RoboCup awards: RoboCup 2004 Engineering Challenge Award, RoboCup 2007 @Home Innovation Award, and RoboCup 2008 @Home Innovation Award. This year UChile HomeBreakers will continue using our social robot, Bender, which obtained the RoboCup @Home Innovation Award in 2007 and 2008, and in 2012 obtained the 6th place in the RoboCup competition. For the 2017 competitions, the main improvements in Bender's hardware and software are: a new head system with an improved face that better expresses emotions, a deep learning system for object recognition, face detection/recognition, the recognition of facial features like age and gender, and improvements in speech recognition using PocketSphinx.

Robocup Articles	03'	04'	05'	06'	07'	08'	09'	10'	11'	12'	13'	14'	15'	16'
Oral	1	2	1	1	2	3	2	2	-	-	1	1	1	-
Poster	1	1	1	-	3	2	-	-	2	1	2	1	3	2

Table 1. UChile's team publications.

3 Current research

Our main technology and scientific contribution is the application of the deep learning paradigm in the Bender robot.

3.1 Deep Learning in Computational Vision for domestic robots (Main research)

Semantic segmentation and place recognition using pre-trained CNN features One of the most complex problems in the area of image processing and computer vision, which has acquired great relevance in recent years, is Semantic Segmentation. This problem consists of determining the label or category for each of the pixels of an image. Semantic Segmentation goes beyond object recognition because all of the pixels must be classified and the borders of the objects must be clearly delimited.

In this context a Semantic Segmentation system was developed. The system (Figure 1) uses CNN pre-trained features, place context information, and the Gpb algorithm (superpixel generation) to improve Semantic Segmentation results. In addition, an edge detector was integrated that allows for improvements in the results of the CNN, with respect to classification by network pixel.

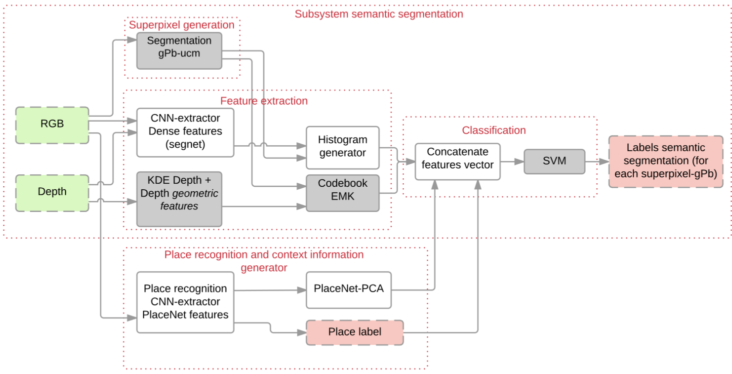

Fig. 1. System diagram of the semantic segmentation and place recognition.

The CNN's pre-trained features performed better than the Hand-crafted methods tested: SIFT-BOW, LBP, 3DLBP. The VGG architecture was the one that gave the best results in place and image recognition using the database NYUv1[1], as can be seen in Table 2.

Network	Accuracy
Alexnet Imagenet (layer3)	75.57%
Alexnet Placenet (layer4)	79.06%
VGG16 Imagenet (layer3)	80.58%
VGG16 Placenet (layer3)	83.76%
GoogleNet Imagenet (layer4)	78.91%
GoogleNet Placenet (layer4)	82.70%

Table 2. Place and image recognition performance.

Region proposal and object classification using CNN with Random Forest and SVM The ability of recognizing objects is of paramount importance in service robotics. Recent approaches used in service and domestic robots are mainly based on analyzing bounding boxes inside the image for detecting objects. As the potential number of bounding boxes to be processed is prohibitively high, algorithms that generate a small number of candidate regions, named region proposals, have been developed. Then proposals are classified by using a Convolutional Neural Network (CNN), enabling accurate and efficient detection even with thousands of object categories.

The emphasis of this research work is to improve the results of the Faster R-CNN system[2] , which is characterized by integrating in a same convolutional architecture, a network that generates region proposals, and another that is able to classify these proposals and predict the location and class of objects within an image. SVM and Random Forest were then used as the final classifiers.

The first system implemented is composed by Faster R-CNN and a random forest classifier. The analysis of the effect of the number of trees in the classifier performance is performed in the validation set of the PASCAL VOC07 database, obtaining the best performance is achieved with 100 and 160 trees in the forest, reaching a 84.6% MAP , while the lowest yield was obtained with 20 trees (MAP = 82.9%).

3.2 Robotic indirect object search

A bayesian based methodology for indirect object search Searching for objects is an important ability required for executing tasks like retrieving requested objects. Currently in the RoboCup@Home, there are predefined manipulation locations, which is limiting as it assumes that objects cannot be placed anywhere. Sometimes the object to be found is hard to detect, but other related objects are easily detectable. Then, detection of secondary objects can help to detect the object to be found if a spatial relation between them is known. This approach is named indirect search.

A novel system that is able to integrate positive and negative detections of secondary objects was proposed and tested. The system stores a single probability map about the presence of the main object, and all positive and negative detections of all objects are used for updating the map. Each detection (positive or negative) generates an observation likelihood with the same size of the map. Then, likelihoods about secondary objects are transformed into likelihoods of the main object by applying a convolution with a spatial relation masks. In addition, different strategies for re-planning based on the current probability map were implemented and compared.

The proposed system is able to achieve state of the art object search. Experiments were performed with the robot Bender, by using a L&R SIFT object detector and an ASUS camera. The task consisted on finding a small object near a larger, easier to detect object in less than 5 minutes. The scene had a size of 6.6[mt] x 9.4[mt], and a configuration similar to rooms used on the RoboCup@Home competition. The objects were placed following a predefined spatial relation. The system uses detected objects as inputs, updates the probability map, and generates new plans for finding the main object in an efficient way. Seven search algorithms were compared on Bender by using 15 trials per algorithm variant. The baseline (uninformed search) has a 40% success rate, informed search using particles [3] has a 73% success, and the proposed informed search method has an 87% success rate.

4 Reusability and applications in the real world

Bender has been conceived as a social robot since his beginnings, designed to be used for the RoboCup@Home league. However, the main idea behind him was to have an open and flexible testing platform that can be used in other application domains. As a social robot, Bender constantly makes public appearances for many and very diverse events and interacts with children and adults alike. He has recently taken part in an opera performance (Orfeo) as one of the characters, visited and entertained children at a hospital, made some demonstrations at different technology fairs across the country and has even made some television appearances as one of the guests or temporary host of the show. To this day, Bender has given talks to more than 3,000 school children. The talks have been given in classrooms and laboratories to groups of 20-25 children, and in a big auditorium to more than 200 children in one session. Also Bender frequently participates in public technology trade fares and events for promoting technology among children and the general public (see pictures in Figure 2).

On the topic of reusability, since Bender was developed to be an open and flexible testing platform that could be used in other applications, almost every piece of software that was developed specifically for the usage on Bender can be used in any other robot with similar characteristics, with only minor changes. In particular, most of the software designed for Bender is under development for usage on the new Pepper robot.

Fig. 2. a) Bender at science exposition. b) Bender at the opera. c) Bender at children's hospital.

5 Conclusions and future work

In the present TDP the main developments of UChile HomeBreakers team for the 2017 RoboCup competitions have been described. As in the last RoboCup the team competed in, the robot that will be accompanying them is Bender, a robot fully designed and developed in the Robotics Laboratory of the Universidad de Chile. In order to succeed in the RoboCup@Home league, Bender has been outfitted with capabilities such as speech recognition, object recognition, face detection and recognition, navigation and obstacle avoidance, map-generation and self-localization, and object manipulation, amongst others. With these abilities and an improved set of hardware, Bender will hopefully have a good demonstration at the 2017 @Home competitions. This year Bender has two important improvements: a new head system including an improved face that shows emotions better, and a deep learning system for object recognition. In addition, improvements were made in speech recognition using PocketSphinx.

For future work, we are conducting research on the following topics:

- A new torso that allows vertical movement for better manipulation reach.
- A prototype for a robotic hand more similar to a humans hand, allowing for a better grip on oddly shaped objects.
- The development of a high-level behavior system for the robot, that uses the stored information in the long-term memory. This will allow the robot to have better interactions with people, since long term relationships can be established.

6 Acknowledgments

This work was partially funded by Fondecyt Project 1130153.

Bibliography

References

1. N. Silberman and R. Fergus. Indoor scene segmentation using a structured light sensor, 2011.
2. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, *Advances in Neural Information Processing Systems 28*, pages 91–99. Curran Associates, Inc., 2015.
3. Alper Aydemir, Kristoffer Sjöö, and Patric Jensfelt. Object search on a mobile robot using relational spatial information. In *Proc. of the 11th Int Conference on Intelligent Autonomous Systems (IAS-11)*, August 2010.
4. Ros website.
5. Davis E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009.
6. Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. Openface: A general-purpose face recognition library with mobile applications. Technical report, CMU-CS-16-118, CMU School of Computer Science, 2016.
7. Gil Levi and Tal Hassner. Age and gender classification using convolutional neural networks. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) workshops*, June 2015.
8. Luca Mella and Daniele Bellavista. Recognizing emotional states in faces.

Robot Bender Hardware Description

We have improved our Bender robot for participating in the RoboCup @Home 2017 competition. The main idea behind its design is to have an open and flexible platform for testing our new developments. We have kept that idea in our improvement. The main hardware components of the robot are (see Figure 1). Specifications are as follows:

Fig. 3. Robot Bender

– Base: The whole robot structure is mounted on a mobile platform. The platform is a Pioneer 3-AT, which has 4 wheels, provides skid-steer mobility, and is connected to 2 Hokuyo URG-04LX lasers for sensing. This platform is endowed with a Hitachi H8S microprocessor. Two notebooks Dell Alienware are placed on the top of the mobile platform with the task of running the navigation and vision modules.
– Chest: The robot's chest incorporates a tablet PC as processing platform; an HP 2760p, powered with a Core i5-2520M Processor (2.50 GHz, 3 MB L3 cache, 2 cores/4 threads) and 4 GB DDR3 PC3-10600 SDRAM (1333 MHz). The tablet includes 802.11bg connectivity. The screen of the tablet PC allows: (i) task of running the speech and manipulation modules and (ii) visualization of relevant information for the user.
– Left and right arms: Both arms are mounted on the torso, each of them has 6 DOF with anthropomorphic configuration. The arms have lightweight design using carbon fiber as main component, this allow to reach a weight about 1 kilogram, enabling safe human interaction applications. The arm actuation is based on Robotis Dynamixel servomotors (3 MX-106, 2 RX-64 and 1 RX-28).
– Grippers: Each arm has a 2 DOF gripper as a tool tip. The grippers use two servomotors (RX-28) as actuators.
– Fieldbus: Robotis Dynamixel servomotors use RS485 based fieldbus, this physical layer allows multidrop comunication and resistance to electromagnetic interference produced in motor drives. In order to reduce cabling and complexity we reuse this fieldbus for custom devices, allowing fast integration of different components as nodes.
– Head: A new 3D printed light-weight head has been developed, with a comprehensive set of new features when compared to the old heavy-weight aluminum head. In order to improve the emotional expressions of the robot, we

added 16 RGB LEDs (WS2812) to each eye, and 4 to each cheek; allowing the robot to cry and to blush. We used two RX-64 Dynamixel motors in pan-tilt configuration, allowing the head movements. The control interface use RS485 Dynamixel based fieldbus.
- 3D Vision. The robot is powered with a Asus Xtion Pro Live over its head (see Figure 1). These device allow for people detection, obstacle detection, at surface detection, and also for object detection while grasping.
- Robot dimensions: height: 1.61m, width: 0.65m and depth: 0.52m.
- Robot weight: 54kg.

Robot's Software Description

The main components of our software architecture take place in one Alienware notebook, while the Navigation and Mapping module run on the second Alienware notebook. Both notebooks use Ubuntu 14.04, and they communicate with each other using ROS Indigo [4]. The Speech, Manipulation and Head modules are running in the HP 2760p are also controlled through ROS.

For our robot we are using the following software:

- Navigation, localization and mapping: ROS is used to implement the functionalities required by the Navigation and Mapping module, among them localization, collision avoidance and logging, map building, and SLAM. For localization, an AMCL particle filter (autonomous monte carlo localization) is used. For navigation, a 2D costs map, with Dikjstra for overall planning and Dynamic Rollout for Local planning, is used.
- Speech recognition: The CMU PocketSphinx library is used to implement a grammar based speech recognition module.
- Speech synthesis: To implement Speech Synthesis the Festival software [13] is used. For speech synthesis a second node is used (from a package ROS called sound play) to make the text to speech synthesis.
- Arms control and two-hand coordination: MoveIt, a ROS package, is used to communicate with the robot to get current state information (positions of the joints, etc.), get point clouds from the robot sensors and to talk to the controllers on the robot.
- Face detection/recognition: dlib [5] and Openface [6].
- Facial features recognition: age and gender recognition [7], emotion recognition [8].
- Object recognition: We use a combination of SIFT LR and VFH. In this hybrid method, the SIFT algorithm is applied it over a subset of the image, obtained by detecting a planar surface in the depth image and then selecting only the blobs that are upper than the plane. If the blobs remain not identified, the VFH algorithm is applied on them. This algorithm is able to use both visual and shape information.