Deep-ChildAR bot: Educational Activities and Safety Care Augmented Reality system with Deep-learning for Preschool

Yoon Jung Park
Media System Lab at Yonsei University
yoonjung.park@msl.yonsei.ac.kr

Hyocheol Ro
Media System Lab at Yonsei University
hyocheol.ro@msl.yonsei.ac.kr

Tack-Don Han
Media System Lab at Yonsei University
hantack@msl.yonsei.ac.kr

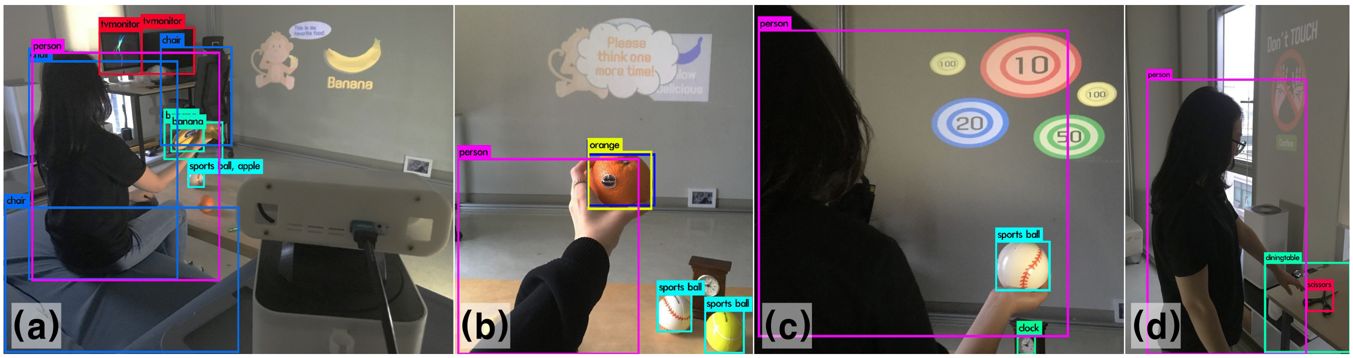

Figure 1: Deep-ChildAR bot scenarios based on Deep learning (a) Our Hardware and Educational activity: Matching game correct (Banana detection). (b) Mathcing game incorrect (Orange detection). (c) Educational activity: Shooting game (Sports ball detection). (d) Safety care function: Augmenting warning information and notifiying guardian (Scissors detection).

ABSTRACT

We propose a projection-based augmented reality (AR) robot system that provides pervasive support for the education and safety of preschoolers via a deep learning framework. This system can utilize real-world objects as metaphors for educational tools by performing object detection based on deep learning in real-time, and it can help recognize the dangers of real-world objects that may pose risks to children. We designed the system in a simple and intuitive way to provide user-friendly interfaces and interactions for children. Children's experiences through the proposed system can improve their physical, cognitive, emotional, and thinking abilities.

CCS CONCEPTS

• **Human-centered computing** → **Mixed / augmented reality**; **Ambient intelligence**; • **Computing methodologies** → **Machine learning**.

KEYWORDS

Projection based Augmented Reality; Educational game; Kid safety; Deep learning; Object detection

ACM Reference Format:

Yoon Jung Park, Hyocheol Ro, and Tack-Don Han. 2019. Deep-ChildAR bot: Educational Activities and Safety Care Augmented Reality system with Deep-learning for Preschool. In *Proceedings of SIGGRAPH '19 Posters.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3306214.3338589

SIGGRAPH '19 Posters, July 28 - August 01, 2019, Los Angeles, CA, USA

ACM ISBN 978-1-4503-6314-3/19/07.
https://doi.org/10.1145/3306214.3338589

1 INTRODUCTION

Augmented reality (AR) technology is an important element education technology that can learn from using various types of information and media data such as videos and images in real environments[Wu et al. 2013]. Because AR technology can integrate real environments and digital learning elements, its educational potential for enhancing contextuality for learning is high. Marker-based AR technology is often utilized in educational AR systems to represent virtual objects in real-world environments via marker. In order to improve the accuracy and robustness of marker recognition, studies are underway that are focused on learning markers using deep learning technology[Dash et al. 2018]. However, it is necessary to purchase or manufacture dedicated tool cards for marker-based AR card games. Because marker-less AR technology can resolve the hassles of attaching markers and offer a more seamless experience to users, it is used for educational applications such as AR pop-up books. The ultimate goal of AR technology is to provide appropriate virtual information to users on the fly by having the technology recognize objects and situations in the real world. Deep learning can allow the system to recognize unknown environments. Thanks to this strength, AR systems aimed at professional education related to medical and military training have been actively developed using deep learning. However, there is a lack of developmental education research aimed at preschoolers. Here, we propose a pervasive robot system for the education and care of preschoolers that utilizes everyday, real-world objects as metaphors for educational tools via deep learning technology. We describe the proposed system and present educational activity and safety care scenarios for children.

Figure 2: Deep-ChildAR Bot Hardware

2 Deep-ChildAR SYSTEM

2.1 Construction AR environment

We reconstruct a three-dimensional (3D) map to provide applications for education and care in an immersive and smooth way to users real environments via projection. We receive color and depth image inputs and then apply 3D point cloud-based map construction technology. We use a feature point matching method[Whelan et al. 2015] that calculates pose by extracting feature points from input images and then compares them with the feature points of the previous frame. We perform the Randomized Hough Transform algorithm based on the segmented areas to detect plane areas. Projection areas are preliminarily selected by minimizing the projection distortion using the normal vector of the maximum area. The optimal projection area is determined by the provided applications and user behavioral intention. The determined optimal projection area is provided to the user by projecting applications and interfaces.

2.2 Object detection using Deep learning

There are many difficulties associated with providing applications by detecting objects present in the real space on the fly that are beyond test-bed and not defined in advance. We attempt to overcome these difficulties by applying deep learning-based object recognition technology. We use YOLO v3[Redmon 2016] (45fps via GTX 1080Ti) with a fast processing time in order to provide users with information, content, and interfaces suitable for real-world objects in real time. We perform object detection in real time via a webcam.

3 APPLICATION

3.1 Scenario design

The designed scenarios can be classified into the following two types: 1) educational activity and 2) a safety care function. First, the educational activity involves improving the children's cognitive abilities to recognize objects by using objects present in the real world as metaphors for educational tools. It augments information such as shadow shapes and sounds from which we can infer objects such as clocks, cell phones, and bottles present in everyday life. Children can solve questions by directly selecting the objects. It is expected that the children will develop muscles as well as cognitive reasoning by touching and interacting with actual objects. Next, a shooting game that is aimed at improving the children's physical abilities dynamically provides random target positions via a pan-tilt system. It can turn a real environment into an AR space for a ball game, thereby letting children naturally engage in physical exercise. The safety care function is for preventing accidents. If the children approach windows, sharp knives, or scissors, this function can prevent accidents by augmenting warning information and notifying the children's guardians immediately via an alarm.

3.2 Interaction design

3.2.1 Spatial touch interaction. It is difficult to build sensors and environments in the real world to determine touch. We implement spatial touch interaction by applying Wilson's method[Wilson 2010]. To do so, we store the depth data of the information-providing area via the depth camera attached to the Deep-ChildAR bot hardware. When users come into contact with an area within the set two thresholds(max, min depth data), this is recognized as a touch.

3.2.2 Object touch interaction. Here, we apply a deep learning framework to utilize real-world objects as educational tools. It was necessary to design and implement an interaction method between real-world objects and the children's hands. The hand was set as the region of interest (ROI) using a two-dimensional (2D) image-based hand detection technique. The overlap between the bounding box (localization of object detection) and the ROI (hand) determines whether the object is selected. In addition, via classification information derived from object detection, it determines whether educational game questions are true or false. Furthermore, the safety care function can be activated using classification information.

4 CONCLUSION AND FUTURE WORK

The proposed system is a portable projection-based AR robot system that can support the education and safety of preschoolers. It performs real-time object detection for real-world objects using a deep learning framework. This system allows real-world objects to be utilized as metaphors for educational tools, which is expected to develop children's cognition, multisensory, body function, etc. It also provides a safety care function for preschoolers because objects present in everyday life can pose risks to children's safety. One limitation of this work is that it is difficult to determine whether there is any contact between hands and objects only with 2D image information (i.e., without any geometry information). We can solve a vision occlusion issue using voice recognition in the future; such work would promote the development of children's auditory and talking abilities. In addition, it is expected that various scenarios for seamless interaction design and multi-sensory development can be derived by fusing deep learning technology for emotion and behavior recognition with face recognition.

ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No.NRF-2018R1A2A1A05078628).

REFERENCES

Ajaya Kumar Dash, Santosh Kumar Behera, Debi Prosad Dogra, and Partha Pratim Roy. 2018. Designing of marker-based augmented reality learning environment for kids using convolutional neural network architecture. *Displays* 55 (2018), 46–54.

Joseph Redmon. 2013–2016. Darknet: Open Source Neural Networks in C. http://pjreddie.com/darknet/.

Thomas Whelan, Michael Kaess, Hordur Johannsson, Maurice Fallon, John J Leonard, and John McDonald. 2015. Real-time large-scale dense RGB-D SLAM with volumetric fusion. *The International Journal of Robotics Research* 34, 4-5 (2015), 598–626.

Andrew D Wilson. 2010. Using a depth camera as a touch sensor. In *ACM international conference on interactive tabletops and surfaces.* ACM, 69–72.

Hsin-Kai Wu, Silvia Wen-Yu Lee, Hsin-Yi Chang, and Jyh-Chong Liang. 2013. Current status, opportunities and challenges of augmented reality in education. *Computers & education* 62 (2013), 41–49.