Deep-Learning-Based Designed Weight Picking Noodle-like Object

Nattapat Koomklang, Sakmongkon Chumkamon, Prem Gamolped, Tomofumi Tsuji, Eiji Hayashi
Department of Mechanical Information Science and Technology, Kyushu Institute of Technology,
680-4 Kawazu, Iizuka, Fukuoka 820-8502, Japan

Abbe Mowshowitz
Department of Computer Science, The City College of New York,
160 Convent Avenue, New York, NY 10031, USA

E-mail: koomklang.nattapat417@mail.kyutech.jp
http://www.kyutech.ac.jp/

Abstract

For food packaging line, manual picking up of the noodle-like objects according to specific weight requires worker's experience for picking quickly and accurately. This article presents a robot arm with a 6-finger gripper picking up the noodle-like objects in specific weight using deep-learning-based to find the best possible action. For measuring the action, we use direct variation to probability of picking action at specific weight in this research use value for the likelihood of weight probability given an action. To find likelihood of weight probably deep-learning-based and use normal distribution for model distribution of the systems. For evaluation we passed any possible action to the network and find action that get maximum likelihood.

Keywords: Deep Convolutional Neural network, Machine Learning,

1. Introduction

Noodle-like objects being picked by robots to a specific weight is challenging for designing a control system. Nowadays, packing noodles into food containers requires experienced workers to pack them up quickly because they can estimate the importance of each pick visually and with experience. This paper proposes a Deep-Learning Neural Network to estimate the robot's grasping action for picking up noodles-like objects from RGB-D camera input with a 6-finguer gripper and cartesian robot movement.

Deep-Learning Neural Network is used in a variety of ways to control robots, especially in picking up things such as dex-net [1] used to predict confidential for grasping object from a parallel-jaw, and vacuum-based suction cup gripper, TossingBot [2] is the network finding the best grasping point for throwing to designed destination. The main concept for the robots grasping network [3], [4], [5], [6], [7] is to find the confidentiality of the state and action of the robot. Commonly confidentiality is represented by the probability of positive action, E.g., successed and failed action. This paper used the likelihood of weight estimation given state (RGB-D image) and robot action to represent the confidence score.

At present, Deep learning has succeeded in the detection, classification, and regression task, especially autoencoder [8], which is the key to representing the latent feature in this research. We use RESNET18 [9] and variational inference [10] for encoding data from RGB-D images. After that, pass encoding data and action of the gripper to fully connected networks for the coefficient of variation in a gaussian distribution.

In this research, the datasets used for training the network are collected from random actions in real environments with fake noodle-like objects with 20,000 data.

Fig. 1 6-fingers gripper and noodle-like objects

2. Methods

2.1. *Model*

Normally for a measure, the confidence score can represent the probability of success. But collecting datasets from the real environment is hard to collect covering all possible action. In this work we represent to use the probability of weight estimation instead because it can predict with regression model. From *Bayes' theorem* we can assume that the probability of weight estimation is direct variations to the probability of success.

$$P(successed|s, x, t_w) \propto P(w|s, x) \qquad (1)$$

Where s is stage in this research is RGB-D image at grasping point at the center of the image. x is picking action and t_w is target picking weight. w is estimation weight.

This research proposed a deep neural network model that predict the picking weight from RGB-D image and action $x = \{x_{gripper\ width}, x_{gripper\ depth}\}$ where $x_{gripper\ width}$ is a distance between the end-tip of gripper. And $x_{gripper\ depth}$ is a distance from reference base to the end-tip of gripper in $z - axis$. The model predicting the weight and its covarion. The model shows in Fig. 2

For RGB-D image, we applied RESNET18 and variational inference for estimate latent parameters in gaussian distribution coefficients. After that we use reparameterization trick (In eq.2) for convert gaussian distribution coefficients to latent parameters (z)

$$z = \mu + \rho\epsilon \qquad (2)$$

Finally, we concatenation latent parameters z and action x through 4-layer fully connected neural networks with leaky relu activation function.

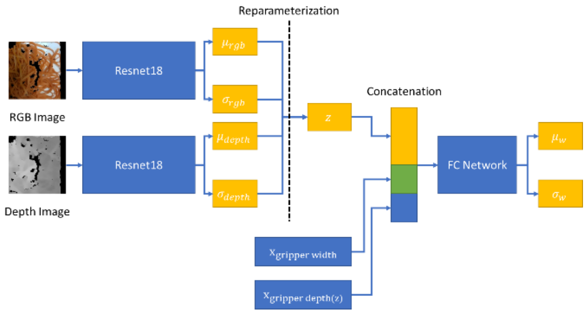

Fig. 2 Weight Estimate Network

2.2. *Loss*

The optimization of the network is the evidence lower bound (ELBO) like in other variational methods as show in eq. Fig. **3**

$$L_{\theta,\emptyset} = \log(p_\theta(w|x)) - \beta D_{KL}(q_\emptyset(z|x)||p(z|x)) (3)$$

The first term, the ELBO term, is a lower bound on the log-likelihood of the data. The second term is the Kullback-Leibler (KL) divergent where β is reduce hyper-parameter.

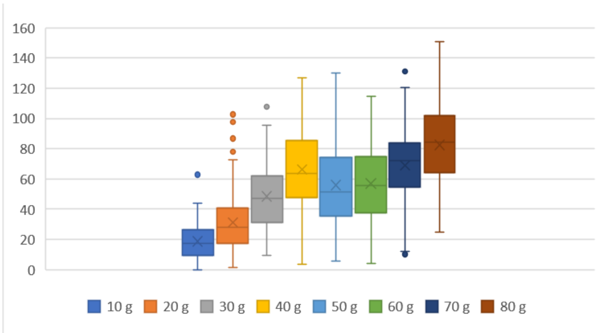

Fig. 3 Picking weight in experimental at 10 – 80 target weight

3. Results

To evaluate the model, we use the network training with 2000 epochs and sampling with possible action to the network between gripper width 30 to 80 mm with step 2.5 mm and gripper depth 0.025 to 0.020 m with step 0.025 m and x, y points every 1 cm grid. After that, we select the action from the maximum likelihood value. The data was collected by setting target weights at 10, 20, 30, 40, 50, 60, 70, 80 g. each target weight accumulated around 100 data points.

Fig. 3 shows the picking weight at each target weight. From the experimental results, it tends to be picked according to the specified weight. But its distribution is relatively high may be due to the sampling for the practical, not enough action point to get the real action.

4. Conclusion

In this paper, we proposed a deep-learning neural network to measure the confidence score for evaluating the action of picking noodle-like with a specified weight. The results of the experiment tend to be picked according to the specified weight. But the distribution is relatively high.

References

[1] J. Mahler, M. Matl, V. Satish, M. Danielczuk, B. DeRose, S. McKinley and K. Goldberg, "Learning ambidextrous robot grasping policies," *Science Robotics,* vol. 4, p. eaau4984, 2019.

[2] A. Zeng, S. Song, J. Lee, A. Rodriguez and T. Funkhouser, *TossingBot: Learning to Throw Arbitrary Objects with Residual Physics,* arXiv, 2019.

[3] H. Cao, G. Chen, Z. Li, J. Lin and A. Knoll, *Lightweight Convolutional Neural Network with Gaussian-based Grasping Representation for Robotic Grasping Detection,* arXiv, 2021.

[4] F.-J. Chu, R. Xu and P. A. Vela, *Real-world Multi-object, Multi-grasp Detection,* arXiv, 2018.

[5] I. Lenz, H. Lee and A. Saxena, *Deep Learning for Detecting Robotic Grasps,* arXiv, 2013.

[6] D. Liu, X. Tao, L. Yuan, Y. Du and M. Cong, "Robotic Objects Detection and Grasping in Clutter Based on Cascaded Deep Convolutional Neural Network," *IEEE Transactions on Instrumentation and Measurement,* vol. 71, pp. 1-10, 2022.

[7] G. Wu, W. Chen, H. Cheng, W. Zuo, D. Zhang and J. You, "Multi-Object Grasping Detection With Hierarchical Feature Fusion," *IEEE Access,* vol. 7, pp. 43884-43894, 2019.

[8] U. Michelucci, *An Introduction to Autoencoders,* arXiv, 2022.

[9] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* 2016.

[10] D. M. Blei, A. Kucukelbir and J. D. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American statistical Association,* vol. 112, p. 859–877, 2017.

[11] D. P. Kingma and M. Welling, "An Introduction to Variational Autoencoders," *Foundations and Trends® in Machine Learning,* vol. 12, p. 307–392, 2019.

Authors Introduction

Nattapat Koomklang

He received master's degree in engineering in 2018, King Mongkut's University of Technology Ladkrabang in Thailand. He is currently a doctor student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Prem Gamolped

He received master's degree in engineering in 2021 from Kyushu Institute of Technology. He is currently a doctor student at Kyushu Institute of Technology and conducts research at Hayashi Laboratory.

Dr. Sakmongkon Chumkamon

Dr. Sakmongkon Chumkamon received Doctor of Engineering degree from Kyushu Institute of Technology in 2017. He was a postdoctoral researcher at Guangdong University of Technology in 2017-2019. Presently he is a postdoctoral researcher in Kyushu Institute of Technology since 2019. His research interests include factory automation robots and social robots.

Prof. Eiji Hayashi

Prof. Eiji Hayashi is a professor in the Department of Intelligent and Control Systems at Kyushu Institute of Technology. He received the Ph.D. (Dr. Eng.) degree from Waseda University in 1996. His research interests include Intelligent mechanics, Mechanical systems and Perceptual information processing. He is a member of The Institute of Electrical and Electronics Engineers (IEEE) and The Japan Society of Mechanical Engineers (JSME).

Prof. Abbe Mowshowitz

Prof. Abbe Mowshowitz received the Ph.D. degree from University of Michigan in 1967. He has been professor of computer science at the City College of New York and member of the doctoral faculty at the Graduate Center of the City University of New York since 1984. His current research interests lie in two areas are organizational and managerial issues in computing, and network science. In addition to teaching and research, He has acted as consultant on the uses and impacts of information technology (especially computer networks) to a wide range of public and private organizations in North America and Europe.