Exit-Pont setting in Edge-Cloud Solution for Object Detection

王 濤† 胡 力博† 渡辺 裕†
Tao Wang† Libo Hu† Hiroshi Watanabe†

榎本 昇平‡ 史 旭‡ 坂本 啓‡ 江田 毅晴‡
Shohei Enomoto‡ Xu Shi‡ Akira Sakamoto‡ Takeharu Eda‡

† 早稲田大学 Waseda University
‡ NTT ソフトウェアイノベーションセンタ NTT Software Innovation Center

1. Introduction

Object recognition can be performed by many cloud vision API services using deep learning. In this case, images are provided to cloud on the Internet. On the other hand, object recognition at an edge becomes possible due to the evolution of calculation power. In general, recognition accuracy achieved at the edge is less than one at cloud. In this paper, we investigate system-level solution for object recognition by combining edge and cloud network.

2. Related Work

Leveraging hybrid computation resources between edge and cloud have been proposed [1][2]. Compression of feature map which should be transferred from edge to cloud has been proposed [3][4]. However, object recognition at the edge is not considered. In other words, there are not exit points at the edge

3. Framework

In the ideal situation, edge network architecture which provides reasonable performance should be a front part of deeper cloud network architecture which achieves sufficient recognition accuracy. Possible typical network is YOLOv2 and modified YOLOv3 for the edge and cloud.

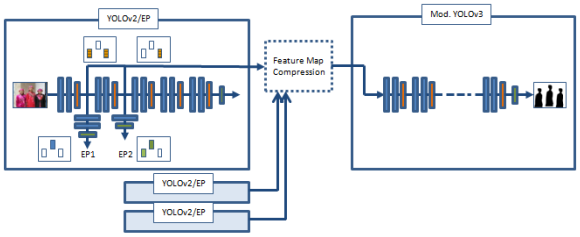

Fig. 1 Example of edge and cloud network combination

4. Edge Operation

For the preliminary performance estimation, how to set and control exit-points at the edge network is investigated. The main branch we use DarkNet19 which is the backbone of the YOLOv2 network. We temporarily set the exit point at the end of 13th convolution layer. In the task of detection, we introduce the class-specific confidence score as a measure of how confident the result we get at exit point. How calculate the confidence score for the branch as shown below.

$$Pr(class_i|object) * Pr(object) * \mathrm{IOU}_{pred}^{truth} = Pr(class_i) * \mathrm{IOU}_{pred}^{truth}$$

5. Investigation

We train and test the network on PASCAL VOC dataset and the platform is NIVIDIA GeForceRTX2070+AMD 3600. Train epoch is 160. As Fig.2 shows, accuracy will smoothly descend with confidence score until score about 0.5 and then the accuracy will drop fast.

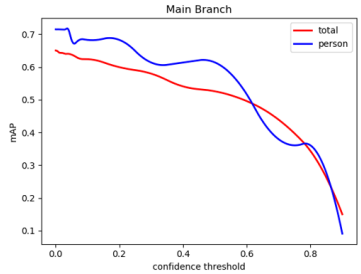

Fig.2 Accuracy of DarkNet19 for varying confidence threshold

When testing on the exit point, we find the detection result is not good enough for practice. Generally, receptive field will influence performance of tasks and larger receptive field will benefit to performance of tasks. Based on Fig.3, we conclude that setting the exit point in the Darknet need larger receptive field and proper size of anchor.

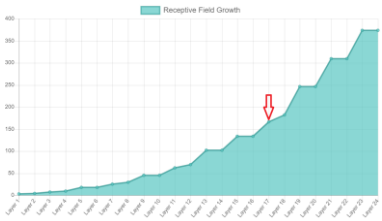

Fig. 3 Receptive filed growth in DarkNet19

6. Conclusion

In this paper, we investigated the performance of edge. We experience on how confidence score influences the accuracy of the main branch result. And find reasons and possible methods for the exit point.

References

[1] S. Teerapittayanon, B. McDanel, and H.T. Kung: “BrancyNet: Fast Inference via Early Exiting from Deep Neural Network,” 2016 International Conference on Pattern Recognition (ICPR2016), pp.2464-2469, Dec. 2016

[2] E. Li, Z. Zhou, and X. Chen: “Edge Intelligence: On-Demand Deep Learning Model Co-Inference with Device-Edge Synergy,” ACM Workshop on Mobile Edge Communications (MECOMM’18), pp.31-26, Aug. 2018

[3] 猪子, 山内, 山下, 藤吉: “クラウド型顔画像解析エンジンにおけるレイテンシを考慮した DCNN の自動分割”, 日本ロボット学会学術講演会 3G1-02, Sep. 2016

[4] H. Choi, and I.V. Bajic: “Deep Feature Compression for Collaborative Object Detection,” IEEE International Conference on Image Processing (ICIP2018) WP. P6.8, Oct. 2018