ACTIVE REINFORCEMENT LEARNING FOR THE SEMANTIC SEGMENTATION OF IMAGES CAPTURED BY MOBILE SENSORS

M. Jodeiri Rad*, C. Armenakis

Geomatics Engineering, GeoICT Lab, Department of Earth and Space Science, Lassonde School of Engineering,
York University, Toronto, Ontario, Canada - {mahyarad@yorku.ca} {armenc@yorku.ca}

Commission II, WG II/6

KEY WORDS: Semantic Segmentation, Active Learning, Reinforcement Learning, Deep Query Network, Deep Neural Network.

ABSTRACT:

In recent years, various Convolutional Neural Networks (CNN) have been used to achieve acceptable performance on semantic segmentation tasks. However, these supervised learning methods require an extensive amount of annotated training data to perform well. Additionally, the model would need to be trained on the same kind of dataset to generalize well for other tasks. Further, commonly real world datasets are usually highly imbalanced. This problem leads to poor performance in the detection of underrepresented classes, which could be the most critical for some applications. The annotation task is time-consuming human labour that creates an obstacle to utilizing supervised learning methods on vision tasks. In this work, we experiment with implementing a reinforced active learning method with a weighted performance metric to reduce human labour while achieving competitive results. A deep Q-network (DQN) is used to find the optimal policy, which would be choosing the most informative regions of the image to be labelled from the unlabelled set. Then, the neural network would be trained with newly labelled data, and its performance would be evaluated. A weighted Intersection over Union (IoU) is used to calculate the rewards for the DQN network. By using weighted IoU, we target to bring more attention to underrepresented classes.

1. INTRODUCTION

Semantic segmentation is the process of building semantic maps, in which the input images are turned into classified raster regions. This field of study is critical to various machine vision tasks like medical image analysis (Gu et al., 2019; Hesamian et al., 2019), autonomous driving (B. Chen et al., 2017), and augmented reality (Guan et al., 2020).

Development of Convolutional neural networks lead into models which could access great performance on the task of semantic segmentation. The Fully Convolutional Network (Shelhamer et al., 2016) was one of the most successful CNN-based segmentation models that fused the output layer with shallower layer's output. Then encoder-decoder networks were introduced, which mapped the low-resolution encoder features to input images feature map (Badrinarayanan et al., 2016). Then U-Net was designed to learn from fewer training data in the biological microscopic imagery case, where there is scarcity of data (Ronneberger et al., 2015). Skip connections had been used on segmentation networks to improve the accuracy and deal with vanishing gradients. FPN model used feature pyramid to better propagate low information into the network (Lin et al., 2017). DeepLab and DeepLabv2 apply several atrous or dilated convolution (Zhao et al., 2017) of the same input with different rates to detect spatial patterns. In DeeplabV3 they created parallel atrous convolution layers, these layers are grouped as Atrous Spatial Pyramid Pooling (ASPP) (L.-C. Chen et al., 2017). Even though these models have achieved state-of-the-art results for semantic segmentation, the need for CNN models to process a considerable amount of high-quality data to generalize effectively and perform robustly is still an open problem.

Preparing training data for semantic segmentation is time-consuming and intense labour, which is an obstacle in taking advantage of abundant data being collected daily with various sensors. Thus, there is a need for a method that could actively learn to decrease the amount of data that needs to be labelled while keeping the same performance.

Another issue with the data for supervised learning methods is that the real world datasets that are used for training, are inherently unbalanced. Naturally, some classes such as sky, vegetation and buildings occupy many more pixels than other classes, while some of these underrepresented classes are much more critical for some applications like self-driving cars. The imbalance is noticeable in the cityscapes dataset (Cordts et al., 2016), which is from street views. In this dataset having 19 classes, the six most underrepresented classes accumulatively occupy less than 2% of the pixels in the training dataset. In contrast, a class like a road occupies more than 36% of the pixels in the training set (Fig. 1). This imbalance reflects the imbalance in the performance of the models.

Active learning is a field of study that addresses the need for a substantial labelled dataset by actively choosing part of the data to be annotated by an "oracle". The approaches had been proven effective in reducing training size while keeping the same performance. Joshi et al., 2009 have developed a method based on uncertainty sampling to perform the image classification task. Later, an adaptive active learning method was proposed for the same task (Li and Guo, 2013). In this method, information density and uncertainty measures were combined with choosing critical instances to be labelled. Then deep reinforcement learning was used as an active learning technique (Fang et al., 2017) for a Natural language processing (NLP) task. Konyushkova et al., 2019 used Deep Q-network (DQN) for active learning for the task of classification. In contrary, the task of semantic segmentation is a computationally complex problem.

---

* Corresponding author
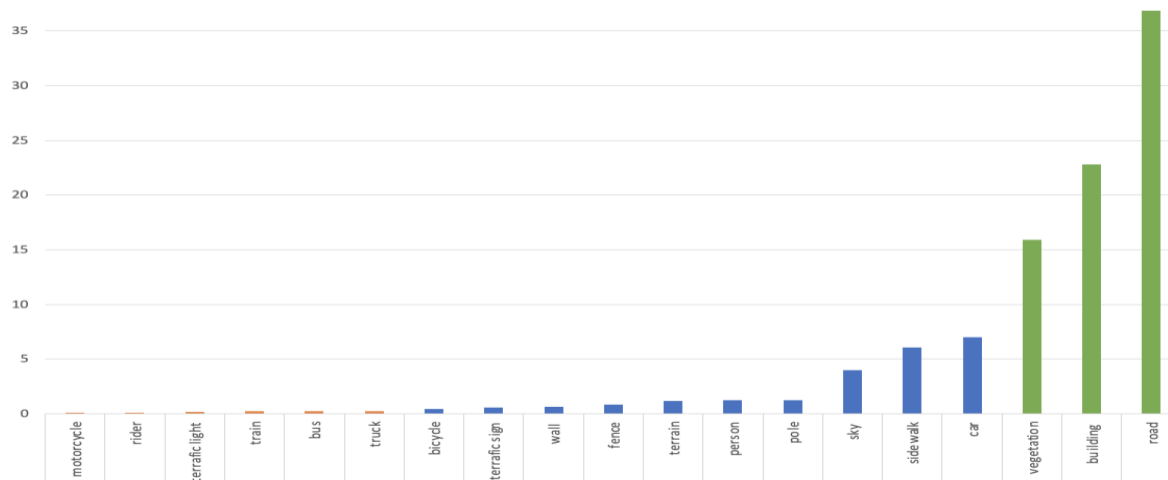
Figure 1. The class distribution of the cityscape dataset

Casanova et al., 2020 handled this problem by applying an end-to-end reinforced active learning method using a DQN as an operational learning strategy to choose regions from the images that would provide the most information. Having the DQN choose from small regions of the image, they address the unbalance dataset issue based on the active learning strategy. To address the problem underrepresented classes is addressed by using mean IoU metric to reward the DQN network.

In this work we aim to address the problem of underrepresented classes even further by using a reinforced active learning method and employing a weighted IoU score for rewarding the DQN network. We hope that by putting more weight on underrepresented classes we can increase the performance of network on these classes even more. We evaluated the performance of this method on the Cityscape dataset and reported the IoU score in comparison to other baselines while using various amounts of data for training.

2. RELATED WORK

There are so many recent works that address the need of the CNN models need for a vast dataset. The encoder-decoders use encoder models that had already been trained for the task of feature extraction. This method is generally known as transfer learning, in which a model would use the knowledge gained while solving a problem and make use of that knowledge for another but related task (Zhuang et al., 2020). These encoders, often called the backbone, help the models generalize well for a complex task while being trained on a smaller dataset. Some well-known backbones architectures are ResNets (He et al., 2015), Xception (Chollet, 2017), and Mobilenet (Howard et al., 2017). Even though transfer learning reduces the need for data severely, a large number of data is still needed for the model to learn the specified task well.

Many semi-supervised and unsupervised learning methods were used to decrease the number of labelled data needed in recent years. Adversarial learning methods have been used to perform the task of semantic segmentation with less labelled data (Hung et al., 2018; Souly et al., 2017). These methods put more effort into finding regions that could be robustly predicted. Li et al., 2020 proposed a way that uses of the mean teacher and student approach to perform the task of semantic segmentation on medical images. Bousias Alexakis and Armenakis, 2021 applied a semi-supervised semantic-segmentation method for change detection.

Some active learning methods address the performance gap between semi-supervised and supervised methods by using active learning methods. In an earlier work done by Vezhnevets et al., 2012, active learning was used to find the most informative nodes of conditional random field (CRF). However, this method depends highly on super-pixels quality. Another work by Nilsson et al., 2020 addressed the problem in source while gathering the data. They used the active learning method to guide the agent to collect the informative data.

Reinforcement learning methods had been used to find the acquisition function for active learning. Ebert et al., 2012 approach to reinforced active learning for classification was based on a Markov decision process (MDP) to create a feedback-driven framework that learns the process during experience without the need for prior knowledge. All these methods only generalize learn the task with less data, dismissing the problems caused by unbalanced classes.

(Kampffmeyer et al., 2016) recognized the problem of imbalanced dataset with remote sensing data and applied various CNNs to evaluate their performance on small objects in the image. (Chen et al., 2019) used a semi-supervised method which utilize maximum square loss instead of minimizing the entropy to prevent them from leaning on to straight forward strategy of choosing easy-to-transfer samples. (Konyushkova et al., 2019) proposed a general purpose data-drive reinforced active learning strategy. Their classification problem is much simpler than the semantic segmentation problem, which increases the computational cost of DQN training. Mackowiak et al., 2018 proposed a strategy where they targeted choosing small regions from images to be labelled by human to maximize the performance of the network while reducing annotation labor.

Casanova et al., 2020 approach to reinforced active learning for semantic segmentation used a data-driven, region-based method that reduces the oracle's labelling effort. This method addressed the class imbalance problem at its core by utilizing a mean intersection over union (MIoU) as a performance metric to evaluate the performance of the segmentation network. The Query network is rewarded for choosing informative regions for training the segmentation network that could improve the MIoU. Additionally, with this region-based approach, the model will have the chance to learn to choose regions from images with the most informative data that the segmentation model had relatively seen less of.
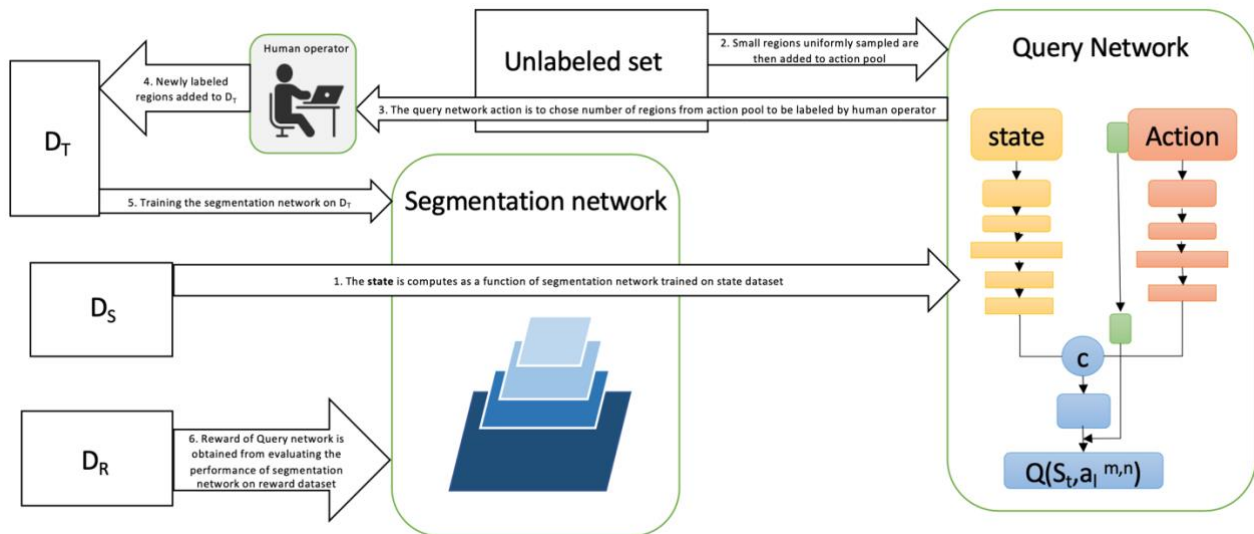
Figure 2. The connections between the query network and segmentation network

3. METHODOLOGY

To increase the attention even further on the underrepresented classes, we propose to use a weighted average of IoU score for computing the reward of the DQN network. By weighing heavier the underrepresented classes while computing the reward, we anticipate that the DQN network would converge to selecting more of data from underrepresented classes, which improving the performance on these specific classes.

In the proposed method we have utilized a Feature Pyramid Network (FPN) (Lin et al., 2017) as a semantic segmentation network and Deep Q-network as the query network. As illustrated in **Figure 2** these two networks are connected to find the optimal policy for the Query network. The optimal policy would be choosing the most informative small regions that had been cropped from images in the training set, in a way that would train the segmentation network to perform well with the least amount of data needed to be labelled. The segmentation model has been pre-trained on a large dataset in order to converge well with less data.

Additionally, by using weighted average IoU, we aim to place additional attention on underrepresented classes, so that the query network would prefer regions with underrepresented classes, which would lead the segmentation network receiving more data from these underrepresented classes in order to perform better. The active learning had been framed as Markov decision process (MDP). The query network is represented as a reinforcement learning agent which would allow the active learning strategy to learn from its own previous experience instead of relying of prior information.

In first step for the data preparation, dataset is parted to four different sets.
1) D_T which would be the training data,
2) D_S is the state set which will be used for computing the state on from the results of segmentation network,
3) The reward set D_R, which will be used to calculate the reward of DQN network from evaluating the segmentation networks performance on this set, and

4) Validation set which will be used to evaluate the performance of segmentation network after it had been trained with the data chosen by the query network.

The reward and state set are chosen carefully from the data that would best represent the whole dataset.

At first the state of DQN network is computed as function of segmentation network on the state set. Then an action pool is being created from regions that had been uniformly selected from unlabeled set. For each action, a sub-action representation has been calculated. Afterward, the query network uses an ε-greedy strategy to select sub-actions from action pool. The ε-greedy is an action selection policy in which the agent takes advantage of prior knowledge by exploitation while at the same time exploring the new options. This approach chooses the action which would prefer highest estimated reward most of the time. The human operator acting as "oracle" annotates the regions that had been chosen by the query network and these annotated regions are added to the training set. The FPN model is trained on the training set and its performance is evaluated based on the reward set. The improvement of the performance of the segmentation network on the reward set in one iteration compared to previous one, provides the query network agent with the reward indicating how well it did on choosing the most informative regions of the images. This loop continues until a predefined budget of labeled data is labeled by operator. Finally, after converging to the optimal policy for query network, the policy is used to choose regions from all the unlabeled set and the segmentation network is trained on with the chosen regions. Finally, its performance is evaluated on a validation set.

For the segmentation network we used the feature pyramid network (Lin et al., 2017). FPN is a fully convolutional feature extractor that takes an image and creates several layers of feature map. The backbone convolutional architectures have no bearing on this procedure. As a result, it serves as a generic approach for constructing feature pyramids inside deep convolutional networks. As stated, this kind of network works especially well on segmenting the smaller objects in the image which usually in the procedure of encoding would become undetectable. The pyramid is built in two ways: from the bottom up and from the top down. In this work a ResNet50 is used as backbone for extracting features (He et al., 2015).

Instead of using the Mean IoU (Equ. 1) metric to penalize the query network, we used a weighted average of IoU across the classes in order to put even more attention on the underrepresented classes. Intersection-Over-Union is a frequently used semantic image segmentation evaluation metric. The prediction confusion matrix (*Figure 3*) is used to compute IoU of each class. In our approach instead of taking the mean of IoU, we computed the W_c as the weight of each class to be inversely proportional to the f_c frequency of that class (Mohapatra et al., 2021).

Figure 3. Confusion Matrix

$$\text{IoU} = \frac{TP}{TP+FP+FN} \quad (1) \qquad Wc = \frac{1}{log(fc+\varepsilon)} \qquad (2)$$

The query network contains two paths, one for state representation and another for action representation. These two paths are concatenated at the end to obtain global features. An entropy map of the regions is computed by using min, max and average pooling operation. Also, two distribution of KL distance between the possible regions, and unlabeled set and unlabeled set is added to action representation. Kullback-Leibler divergence, or simply, the KL divergence (also called relative entropy), quantifies how one probability is different from the other. State representation is completed by adding class distribution features.

To find the optimal policy mapping each state to an action that maximizes the expected sum of future rewards we have used a DQN (Mnih et al., 2013). The DQN is trained by minimizing the loss based on temporal difference (TD) error (van Hasselt et al., n.d.) where we are considering SARSA transitions. The SARSA (State-Action-Reward-State-Action) transition which is defined as $(s_t, a_t, r_t, s_{t+1}, a_{t+1})$ is an iterative algorithm for finding the optimal policy. In this case, the action is defined as selecting number of regions to be labeled by annotators. Then the network is trained using selected regions and the reward is afterwards calculated in a held-out portion of data. To stabilize the training, we use double DQN formulation in which action selection is decoupled from evaluation.

4. EXPERIMENTS AND RESULTS

4.1 Dataset

We use the Cityscape dataset to evaluate the performance of the network (van Hasselt et al., n.d.). The dataset contains training set of 2975 RGB images with the size of 2048×1024 pixels. There are two versions of dataset with 19 and 35 classes, respectively; we used the 19 classes version. We chose 10 images for state set, 200 images for reward set, and 200 images for reward set and 150 images for training the query network. The rest of 2615 images considered unlabelled and are used in D_v, to evaluate the performance of the acquisition function. In other words, 12% of labelled data is used just to train the query network to obtain the optimal policy. In this process each image in the D_T is split to 128 small regions with size of 128×128 pixels to create the action pool. In each step of training the Query network chooses 256 regions to be labelled. The DQN network is trained with budget of 3840 regions.

4.2 Pre-training the segmentation network

The FPN model is first pre-trained on the GTAV dataset (Richter et al., 2016). This dataset contains 24966 synthetic images which have been rendered using Grand Theft Auto 5 video game from car perspective in street scenery. There are 19 semantic classes in the synthetic dataset which makes it compatible with cityscape dataset. Then the model is finetuned on D_T dataset. This pre-training before the labelling process seems to be time consuming, however, it is a small amount of time compared to the time needed for labelling the whole dataset.

4.3 Results

We experimented with 5 sets of labelling budget for the query network which varies in the range from 1% of dataset to 5% of dataset. The performance of segmentation model is evaluated by mean IoU score. Figure 4 presents the results we obtained from evaluating the performance of our segmentation model on the validation set. The segmentation model had been trained on the labelled regions that had been chosen by the query network after it had already been trained on 12% of the dataset that we separated at the beginning. We compared our results with the three baseline results: i) Random which is the uniform random sampling of regions from action pool, ii) the Entropy method which is an uncertainty sampling method which applies the policy of choosing maximum pixel-wise Shannon entropy, and iii) BALD method which select the regions based on maximum cumulative pixel-wise BALD metric (Gal et al., 2017).

Figure 4. Valid mean IoU score of methods compared across different budgets of labeled data.

As presented on **Figure 4**, the proposed RAlis method (Reinforcement Active Learning for image segmentation) both with mean IoU (RALis-MIoU) and weighted IoU (RALis-WIoU) perform better than other baselines. Contrary to our expectation the weighted IoU method results are quiet similar to results obtained by the mean IoU. Investigating further, we noticed that since the reward set is also the representation of the whole dataset, it is also severely imbalanced. Therefore, since the reward is computed as performance of the segmentation network on the reward set, the change of the weighted IoU does not lead to any sensible improvement using this method. Therefore, the effect of the reward set on learning process of Q-network should be further investigated in future work.
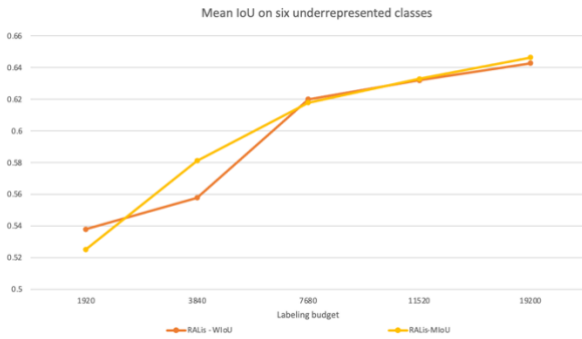
Figure 5. Mean IoU on six underrepresented classes

Figure 5 compares the performance of original RALis model with weighted IoU rewarding RALis model across various budgets of labelled data on six most underrepresented classes: motorcycle, rider, traffic light, train, bus and truck. The performance is very close to one another.
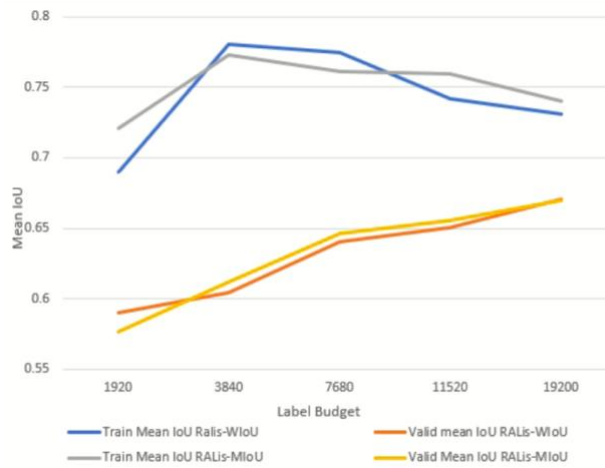

Figure 6. Training and Validation IoU

The training and validation IoU for RALis optimized by mean IoU and RALis optimized by weighted IoU are presented in **Figure 6**. The graph represents the learning behaviour of the reinforced active learning approach. In the smaller sample sizes, the gap between training and validation IoU is large, which suggests overfitting. As the number of budget increase, the gap shrinks to the point that validation IoU become even larger than the Train IoU, which suggests that the more budget the segmentation network has, the better It generalize.
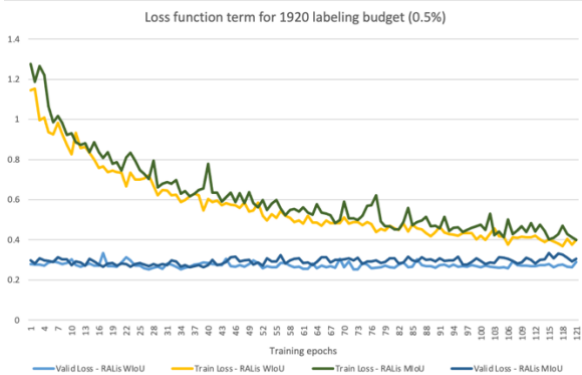

Figure 7. Training and Validation loss with 0.5% labeling budge

Figure 7 graphs the Loss curves for RALis with mean and weighted IoU methods can help us to closely examine the learning behaviour of segmentation model. We can see that all terms decrease - as they were supposed to - during training. The consistent decrease of validation loss displays that the method had not overfitted even in the smallest labelling budget . Also, **Figure 8** Display the training and validation loss in 5% labelling budget. The learning behaviours of both methods are reasonable.
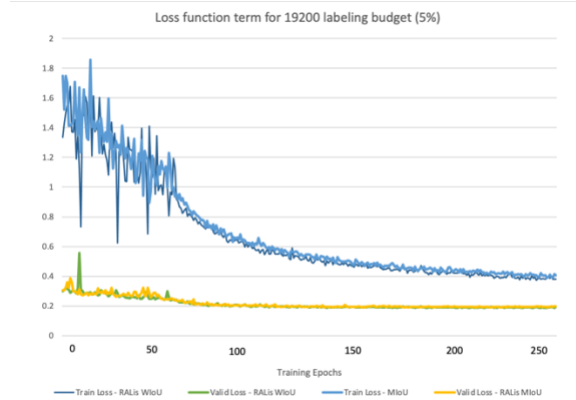

Figure 8. Training and Validation loss with 5% labeling budget

5. CONCLUSION

We used a data-driven region-based reinforcement active learning method for segmenting the data captured by mobile sensor on street scenery. A reinforcement learning agent performed the active learning process by converging to a policy based on ε-greedy approach. The ideal policy would choose the most informative samples from pool of small regions cropped from the image. The goal was to decrease the labor of annotation while keeping the same level of performance.

State and action representation are defined in to be aware of the classes distribution. Additionally, by choosing the weighted IoU for calculating the reward for DQN network, we further increase the attention on underrepresented classes. By increasing the attention on underrepresented classes, the segmentation network results are better not only in general, but also on the underrepresented classes compared to other baseline methods.

Even though, the proposed weighted IoU method did not achieve any noticeable improvement in results compared to the mean IoU approach, it casted a light on importance of reward and state set on this method. We plan to further study the effect of these sets on the training of the Deep Q-network. Additionally, we plan to use different segmentation models to study the effects of segmentation models on success of Q-network.

Another important problem to be further studied, is the performance of this method on other type of data. As it had been mentioned before that this active learning method is data-driven. Therefore, we plan on studying the efficiency of this method on a dataset retrieved by UAV sensors. UAV datasets are usually even more imbalanced than street scenery, therefore, we look forward for applying this method on UAV image set.

REFERENCES

Badrinarayanan, V., Kendall, A., Cipolla, R., 2016. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *ArXiv151100561 Cs*.

Bousias Alexakis, E., Armenakis, C., 2021. Evaluation f Semi-Supervised Learning for CNN-Based Change Detection. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLIII-B3-2021, 829–836.* https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-829-2021.

Casanova, A., Pinheiro, P.O., Rostamzadeh, N., Pal, C.J., 2020. Reinforced Active Learning for Image Segmentation 17.

Chen, B., Gong, C., Yang, J., 2017. Importance-Aware Semantic Segmentation for Autonomous Driving System, *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence.* Melbourne, Australia, pp. 1504–1510. https://doi.org/10.24963/ijcai.2017/208

Chen, M., Xue, H., Cai, D., 2019. Domain Adaptation for Semantic Segmentation with Maximum Squares Loss, *2019 IEEE/CVF International Conference on Computer Vision (ICCV).* pp. 2090–2099. https://doi.org/10.1109/ICCV.2019.00218

Chollet, F., 2017. Xception: Deep Learning with Depthwise Separable Convolutions. *ArXiv161002357 Cs*.

Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. *ArXiv160401685 Cs*.

Ebert, S., Fritz, M., Schiele, B., 2012. RALF: A reinforced active learning formulation for object class recognition, *2012 IEEE Conference on Computer Vision and Pattern Recognition. (CVPR),* IEEE, Providence, RI, pp. 3626–3633. https://doi.org/10.1109/CVPR.2012.6248108

Fang, M., Li, Y., Cohn, T., 2017. Learning how to Active Learn: A Deep Reinforcement Learning Approach. *ArXiv170802383 Cs*.

Gal, Y., Islam, R., Ghahramani, Z., 2017. Deep Bayesian Active Learning with Image Data. *ArXiv170302910 Cs Stat*.

Gu, Z., Cheng, J., Fu, H., Zhou, K., Hao, H., Zhao, Y., Zhang, T., Gao, S., Liu, J., 2019. CE-Net: Context Encoder Network for 2D Medical Image Segmentation. *IEEE Trans. Med. Imaging 38, 2281–2292.* https://doi.org/10.1109/TMI.2019.2903562

Guan, P., Cao, Z., Chen, E., Liang, S., Tan, M., Yu, J., 2020. A real-time semantic visual SLAM approach with points and objects. *Int. J. Adv. Robot. Syst. 17*, 172988142090544. https://doi.org/10.1177/1729881420905443

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep Residual Learning for Image Recognition. *ArXiv151203385 Cs*.

Hesamian, M.H., Jia, W., He, X., Kennedy, P., 2019. Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *J. Digit. Imaging 32, 582–596.* https://doi.org/10.1007/s10278-019-00227-x

Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H., 2017. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *ArXiv170404861 Cs*.

Hung, W.-C., Tsai, Y.-H., Liou, Y.-T., Lin, Y.-Y., Yang, M.-H., 2018. Adversarial Learning for Semi-Supervised Semantic Segmentation. *ArXiv180207934 Cs*.

Joshi, A.J., Porikli, F., Papanikolopoulos, N., 2009. Multi-class active learning for image classification, *2009 IEEE Conference on Computer Vision and Pattern Recognition.* (CVPR Workshops), IEEE, Miami, FL, pp. 2372–2379. https://doi.org/10.1109/CVPR.2009.5206627

Kampffmeyer, M., Salberg, A.-B., Jenssen, R., 2016. Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks, *2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW).* Las Vegas, NV, USA, pp. 680–688. https://doi.org/10.1109/CVPRW.2016.90

Konyushkova, K., Sznitman, R., Fua, P., 2019. Discovering General-Purpose Active Learning Strategies. *ArXiv181004114 Cs Stat*.

Li, X., Guo, Y., 2013. Adaptive Active Learning for Image Classification, in: 2013 IEEE Conference on Computer Vision and Pattern Recognition. *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Portland, OR, USA, pp. 859–866. https://doi.org/10.1109/CVPR.2013.116

Li, X., Yu, L., Chen, H., Fu, C.-W., Xing, L., Heng, P.-A., 2020. Transformation Consistent Self-ensembling Model for Semi-supervised Medical Image Segmentation. *ArXiv190300348 Cs*.

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., Belongie, S., 2017. Feature Pyramid Networks for Object Detection. *ArXiv161203144 Cs*.

Mackowiak, R., Lenz, P., Ghori, O., Diego, F., Lange, O., Rother, C., 2018. CEREALS - Cost-Effective REgion-based Active Learning for Semantic Segmentation. *ArXiv181009726 Cs*.

Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D., Riedmiller, M., 2013. Playing Atari with Deep Reinforcement Learning. *ArXiv13125602 Cs*.

Mohapatra, S., Yogamani, S., Gotzig, H., Milz, S., Mader, P., 2021. BEVDetNet: Bird's Eye View LiDAR Point Cloud based Real-time 3D Object Detection for Autonomous Driving. *ArXiv210410780 Cs*.

Richter, S.R., Vineet, V., Roth, S., Koltun, V., 2016. Playing for Data: Ground Truth from Computer Games. *ArXiv160802192 Cs*

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. *ArXiv150504597 Cs*.

Shelhamer, E., Long, J., Darrell, T., 2016. Fully Convolutional Networks for Semantic Segmentation. ArXiv160506211 Cs.

Souly, N., Spampinato, C., Shah, M., 2017. Semi Supervised Semantic Segmentation Using Generative Adversarial Network, *2017 IEEE International Conference on Computer Vision (ICCV).* Venice, pp. 5689–5697.
https://doi.org/10.1109/ICCV.2017.606

van Hasselt, H., Guez, A., Silver, D., n.d. Deep Reinforcement Learning with Double Q-Learning 7.

Vezhnevets, A., Buhmann, J.M., Ferrari, V., 2012. Active learning for semantic segmentation with expected change, *2012 IEEE Conference on Computer Vision and Pattern Recognition.* Providence, RI, pp. 3162–3169.
https://doi.org/10.1109/CVPR.2012.6248050

Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid Scene Parsing Network. ArXiv161201105 Cs.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., He, Q., 2020. A Comprehensive Survey on Transfer Learning. *ArXiv191102685 Cs Stat.*