CONFIDENCE GUIDED SEMI-SUPERVISED LEARNING IN LAND COVER CLASSIFICATION

Wanli Ma, Oktay Karakuş, Paul L. Rosin

School of Computer Science and Informatics, Cardiff University, Cardiff CF24 4AG, UK

1. INTRODUCTION

Semi-supervised learning approaches help reduce the cost of manual labelling by exploiting a large quantity of unlabelled data. Especially in the application of land cover classification, pixel-level manual labelling in large-scale imagery is labour-intensive and expensive. Previous work on pseudo-label-based semi-supervised land cover classification or semantic segmentation suffers from the low accuracy of pseudo labels and usually shows unsatisfactory performance compared to supervised learning approaches. Also, semi-supervised learning semantic segmentation approaches [1, 2] in computer vision are not always efficient in land cover classification, due to remote sensing data complexities such as heavy class imbalance, the high similarity between some classes (e.g. wetlands and forests), and low resolution.

The existing semi-supervised learning methods pay limited attention to the quality of pseudo-labels when using them to supervise the network. That is, however, one of the key factors that determine network performance. In order to fill this gap, we develop a *confidence-guided semi-supervised learning* (CGSSL) approach to make use of high-confidence pseudo labels and reduce the negative effect of low-confidence labels on training the land cover classification network.

Specifically, the contributions of this work are as follows: (1) we propose a confidence-aware cross-entropy (CCE) loss for semi-supervised land cover classification which is flexible and can be easily transferred to other computer vision tasks such as semantic segmentation. (2) an adaptive mechanism is designed to adjust the threshold automatically (no-manual setting) for judging the quality of the pseudo labels based on their confidence. (3) we promote the investigation of multiple network outputs in terms of an information theory aspect – entropy – to weight confidence levels of pseudo labels from each network. Then, we use this confidential information to optimise the unsupervised training process in semi-supervised land cover classification.

2. METHOD

The framework of the proposed method is shown in Figure 1. In each training iteration, both labelled and unlabelled data are given as input into the networks. The labelled data is used in a regular supervised learning pattern to train these models, where the supervised loss of each network is calculated by the cross-entropy loss function. On the other hand, unlabeled data is utilized to generate pseudo labels, which are exploited to inform each network. Specifically, both supervised and unsupervised learning stages implement the three different networks (PSPNet [3], UNet [4], SegNet [5]), and their weights are shared between the two stages. The prediction of each network is the class probabilities for all pixels of the corresponding input image, and then the predictions from three networks are added linearly after a softmax layer to generate a comprehensive prediction. In this case, if the confidence distributions among the three networks are identical, the operation of linear addition will promote the distribution to be sharp (low uncertainty). Otherwise, the combined prediction will become smooth (high uncertainty). Considering the fact that the information entropy is a measure of uncertainty, we calculate the entropy of classification distribution based on the combined prediction to confirm confidence in the predictions. Furthermore, the proposed CCE loss function is designed to limit the negative contribution of the pseudo labels with high entropy (high uncertainty) to network parameter optimisation.

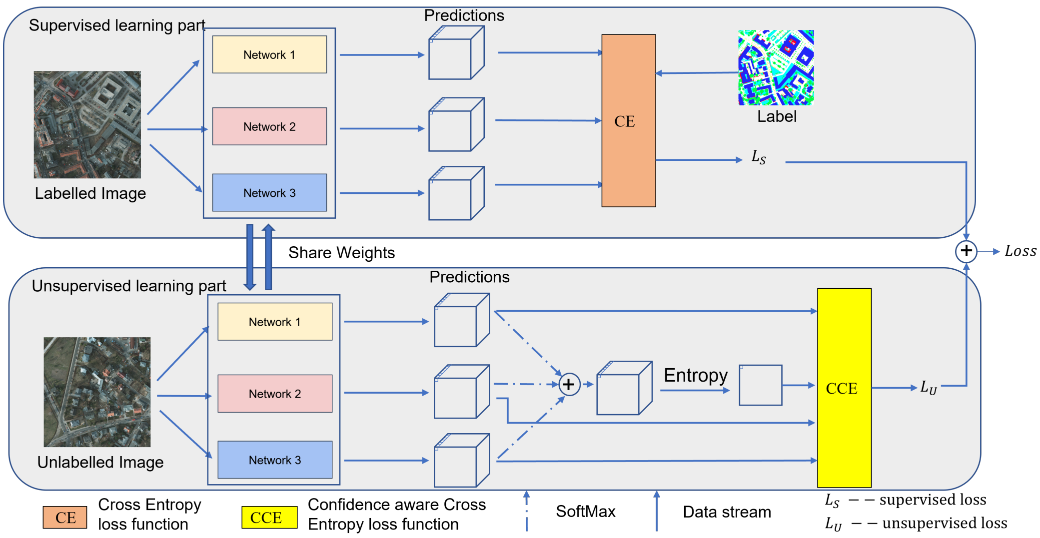


Fig. 1: Overall framework of the confidence guided semi-supervised learning (CGSSL) approach

As shown in Figure 2, the proposed confidence-aware cross entropy (CCE) loss module is used to calculate the unsupervised loss. The aim of CCE loss is to keep the cross entropy loss between the high-quality labels and predictions identical whilst reducing the effect of the low-quality labels on the unsupervised loss at the pixel level. The average value of entropy is regarded as a threshold to divide high-quality and low-quality labels, where the threshold is therefore adaptive to the input entropy and can be adjusted automatically. The prediction with high confidence will be identical, whereas the prediction with low confidence will be multiplied by a small value ranging between 0 and 1 to reduce its effect when calculating the unsupervised loss. Inspired by ([2], [6]), the unsupervised loss is acquired by cross-supervision between predictions from different networks.

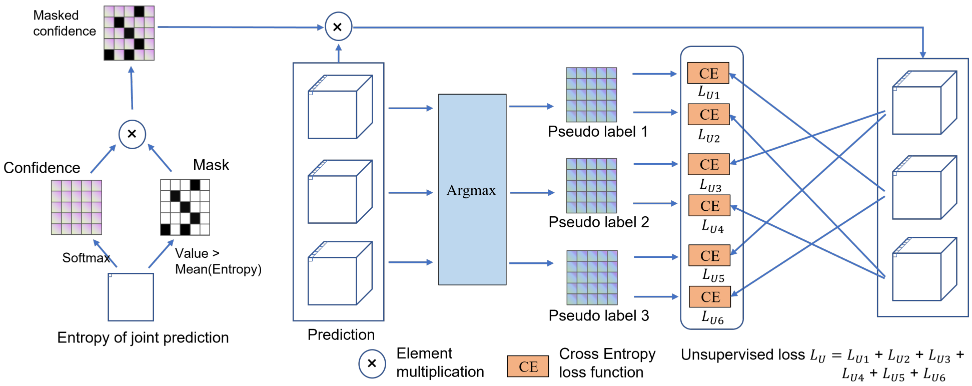
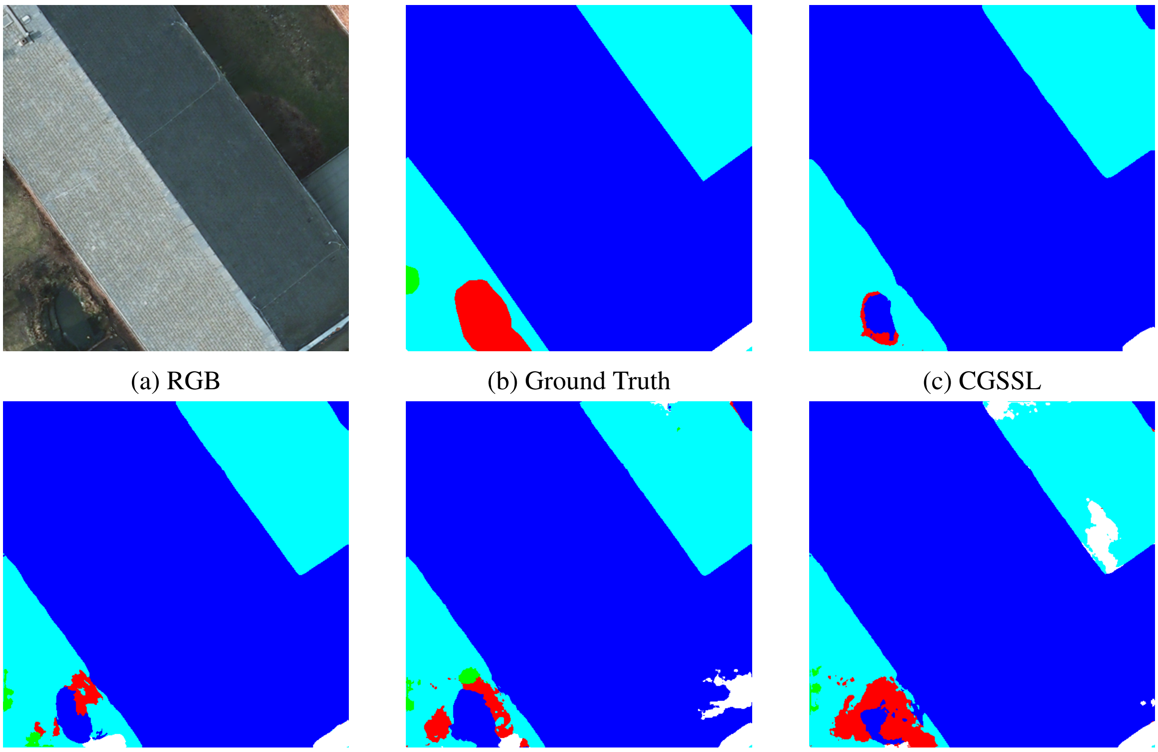


Fig. 2: Confidence-aware Cross Entropy (CCE) module.



(a) RGB (b) Ground Truth (c) CGSSL

(d) CPS (e) Half Labeled Supervision (f) Full Labeled Supervision

Fig. 3: Qualitative results on Potsdam validation set.

3. RESULTS & DISCUSSIONS

We use Potsdam [7] as a test dataset for the original UNet [4], Mean Teacher [1], and CPS [2] with UNet. The number of labelled data used in these semi-supervised learning approaches is only half of the whole Potsdam dataset. We remove the labels of the remaining half and used them in the unsupervised part. We also provide the performance of UNet only in supervised learning patterns for half and the whole labelled data. The same test set is used to evaluate all modules. As shown in Table 1, CGSSL shows the best performance in terms of all performance metrics. Especially, CGSSL improves the precision significantly due to the great reduction of false positives in prediction. Even though CGSSL only uses half-labelled data and unlabelled data, its performance is even better than UNet which is optimised with the whole labelled data.

Table 1: Performance comparison

	Precision	Recall	mIoU	F1	Accuracy
UNet with Full Labelled Data [4]	80.16%	78.07%	68.17%	79.10%	85.59%
UNet with Half Labelled Data [4]	79.00%	76.35%	66.25%	77.65%	84.24%
Mean Teacher [1]	80.88%	77.13%	67.31%	78.96%	84.71%
CPS [2]	82.39%	77.24%	68.54%	79.73%	85.99%
CGSSL	**84.04%**	**78.02%**	**69.48%**	**80.92%**	**86.58%**

In conclusion, we proposed a novel, confidence-guided, cross-entropy loss-based semi-supervised learning for land cover classification applications. Especially, an adaptive loss is provided for the semi-supervised learning to exploit pseudo labels in an information theory perspective, which is also flexible to be transferred to various other semi-supervised learning tasks. The proposed method shows considerable performance and benefit from unlabeled data for land cover classification.

4. REFERENCES

[1] Antti Tarvainen and Harri Valpola, "Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results," *Advances in neural information processing systems*, vol. 30, 2017.

[2] Xiaokang Chen, Yuhui Yuan, Gang Zeng, and Jingdong Wang, "Semi-supervised semantic segmentation with cross pseudo supervision," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 2613–2622.

[3] Hengshuang Zhao, Jianping Shi, Xiaojuan Qi, Xiaogang Wang, and Jiaya Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[5] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[6] Xiangde Luo, Minhao Hu, Tao Song, Guotai Wang, and Shaoting Zhang, "Semi-supervised medical image segmentation via cross teaching between CNN and transformer," in *International Conference on Medical Imaging with Deep Learning*. PMLR, 2022, pp. 820–833.

[7] Franz Rottensteiner, Gunho Sohn, Jaewook Jung, Markus Gerke, Caroline Baillard, Sebastien Benitez, and Uwe Breitkopf, "The isprs benchmark on urban object classification and 3d building reconstruction," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences I-3 (2012), Nr. 1*, vol. 1, no. 1, pp. 293–298, 2012.