
Adaptive-Labeling for Enhancing Remote Sensing Cloud Understanding

Jay Gala¹, Sauradip Nag², Huichou Huang³, Ruirui Liu⁴, and Xiatian Zhu²

¹NMIMS University, ²University of Surrey, ³City University of Hong Kong, ⁴Brunel University London

Abstract

Cloud analysis is a critical component of weather and climate science, impacting various sectors like disaster management. However, achieving fine-grained cloud analysis, such as cloud segmentation, in remote sensing remains challenging due to the inherent difficulties in obtaining accurate labels, leading to significant labeling errors in training data. Existing methods often assume the availability of reliable segmentation annotations, limiting their overall performance. To address this inherent limitation, we introduce an innovative model-agnostic *Cloud Adaptive-Labeling (CAL)* approach, which operates iteratively to enhance the quality of training data annotations and consequently improve the performance of the learned model. Our methodology commences by training a cloud segmentation model using the original annotations. Subsequently, it introduces a *trainable pixel intensity threshold* for adaptively labeling the cloud training images on the fly. The newly generated labels are then employed to fine-tune the model. Extensive experiments conducted on multiple standard cloud segmentation benchmarks demonstrate the effectiveness of our approach in significantly boosting the performance of existing segmentation models. Our CAL method establishes new state-of-the-art results when compared to a wide array of existing alternatives.

1 Introduction

Cloud understanding is indicative and critical in climate science, with significant application in various sectors like renewable energy, disaster management, and environmental monitoring [1]. A viable and scalable approach to this problem is leveraging remote sensing imagery data to examine the patterns and characteristics of clouds over space and time. To that end, cloud segmentation is one of the most fine-grained methods.

Traditionally, cloud segmentation primarily employed threshold-based techniques, hinging on pre-existing knowledge and the distinction between Earth’s surface and clouds [2]. Nevertheless, these methods exhibited vulnerability to intricate backgrounds, thereby constraining their efficacy. Additionally, the selection of the threshold in such methods typically entailed subjective human judgment and expertise, resulting in performance disparities [3].

Recently, deep learning approaches have emerged as promising alternatives to threshold-based methods. They excel in capturing intricate patterns and leveraging extensive datasets, resulting in significant enhancements in cloud segmentation accuracy [4]. Deep learning mitigates the reliance on subjective thresholding while adapting effectively to diverse atmospheric conditions and complex cloud formations [5].

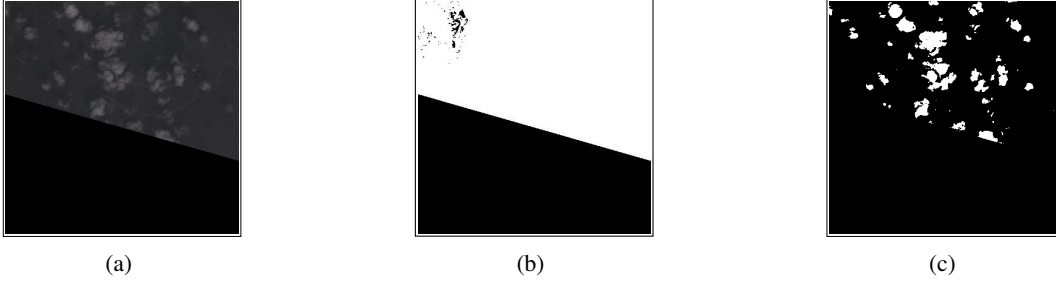


Figure 1: Illustration of noisy cloud mask labeling. (a) A random training image; (b) The original cloud mask label; (c) The mask label obtained using our proposed method.

Nonetheless, all the previous methods typically assume the availability of accurate mask annotations of cloud, which is often largely invalid (see Figure 1 (a)). Making this assumption could generally lead to degraded model performance, but it is still understudied in the literature.

In this paper, we propose an innovative model-agnostic *Cloud Adaptive-Labeling (CAL)* method to address the mentioned fundamental challenge. Our approach starts by training a cloud segmentation model using the original noisy training data. Subsequently, it introduces a *trainable pixel intensity threshold* for adaptively labeling the training images on the fly. The newly generated labels are then employed to fine-tune the model. Importantly, our pixel intensity threshold is dynamically adjusted through an intuitive self-regulating mechanism in response to the evolving training loss: it increases when the loss decreases and decreases when it increases. This objective-oriented thresholding strategy challenges the conventional notion of a single optimal threshold and obviates the necessity for prior threshold estimation. Extensive experiments demonstrate that our method iteratively enhances the quality of training data annotations, leading to improvements in model performance and achieving the new state of the art on a standard cloud segmentation benchmark.

2 Related work

Existing cloud segmentation techniques are categorized into three groups: Threshold-based, Machine learning-based, and Deep learning-based. Threshold-based methods rely on differences in scattering intensity between clouds and atmospheric particles in specific spectral bands [6]. The spectral threshold approach is effective when combined with multi-band features like radiance, reflectance, and normalized vegetation index [7, 8, 9]. These methods work well only when clouds have high visible reflection and cooler infrared temperatures compared to the Earth’s surface. However, their performance depends on prior knowledge and the contrast between clouds and the Earth’s surface. Despite their effectiveness, threshold-based methods struggle to distinguish clouds in the presence of atmospheric aerosols, dust, smoke, and hazy pixels caused by small-scale cumulus clouds or cloud edges [3].

To address the limitations of threshold-based approaches, following-up works employed machine learning algorithms, such as Support Vector Machines (SVMs) and random forests [10, 11], leveraging texture features extracted from remote sensing images as input [12, 13]. However, these methods heavily depended on manually crafted features. Recent advancements have witnessed the adoption of deep learning techniques in cloud segmentation, with subsequent efforts focused on enhancing network architecture and deep feature extraction [14, 15, 16]. For instance, Yao et al. [17] introduced attention modules and boundary refinement blocks to comprehensively capture multi-scale context.

While deep learning-based cloud detection methods exhibit significant advantages over threshold-based and traditional machine learning approaches, their outstanding performance hinges on the availability of extensive, high-quality pixel-level cloud masks [18]. Nevertheless, the creation of such datasets is a challenging and time-consuming endeavor, often resulting in the presence of substantial noisy labels that can introduce uncertainties and negatives into research outcomes (see Figure 1). Interestingly, this problem has never been systematically considered and investigated in the literature. This study marks the inaugural endeavor to address the challenge of noisy cloud mask labeling by introducing a model-agnostic CAL method within the deep learning domain.

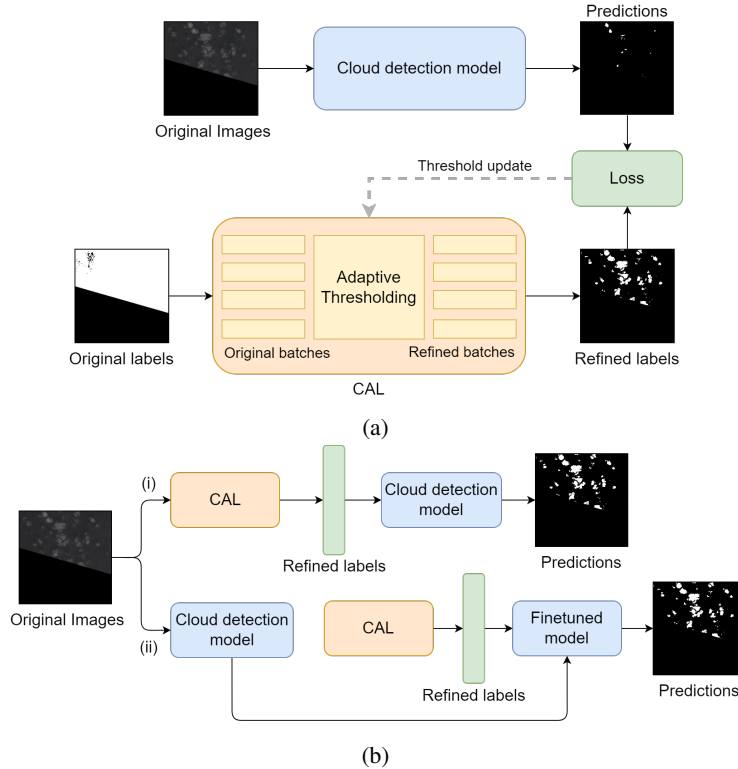


Figure 2: (a) **Overview for our proposed *Cloud Adaptive-Labeling (CAL)* method.** Our approach begins with training a cloud segmentation model using the original training data, which inherently contains noise. We then incorporate a trainable pixel intensity threshold to adaptively label the cloud mask by applying this threshold to the input training images. These newly generated cloud mask labels are subsequently used for fine-tuning the model. Importantly, our threshold is dynamically adjusted based on the training loss: it increases when the loss decreases and decreases when the loss increases. This innovative training objective-driven thresholding strategy not only challenges the traditional assumption of a single optimal threshold but also eliminates the need for prior estimation of an appropriate threshold. (b) **Integrating CAL** (i) during finetuning with existing labels (ii) after finetuning with existing labels

3 Cloud Adaptive Labeling

Overview. Unlike previous methods, which prioritized improving model efficiency and capturing image-label relationships, we propose a method to reduce noise in cloud segmentation datasets. Our *Cloud Adaptive-Labeling (CAL)* is an efficient module that enhances existing cloud detection approaches. CAL is typically plugged in as a post-processing block to relabel existing noisy cloud labels into a refined version. When integrated into the training scheme, CAL boosts the performance of existing pretrained models. As shown in Figure 2 (a), CAL mainly consists of a novel dynamic thresholding module that mitigates the labeling noises.

Problem statement. Given a satellite image I , cloud detection predicts a pixel-level binary classification of whether each pixel is a cloud, outputting a binary mask M . Formally, a training dataset D_{train} contains image-mask pairs (I, M) used to train a cloud detection model, which is then evaluated on a testing split D_{test} .

CAL algorithm. Serving as a model-agnostic plug-in, our CAL is designed to gradually improve mask labels during training, boosting model performance. We initialize a per-mini-batch threshold parameter p_{th} , which learns based on changes in model predictions. We use p_{th} to binarize the input images as $M_{cal} = \phi(I, p_{th})$ where $\phi(\cdot)$ is the morphological image binarization operation. We replace the original/previous mask labels with M_{cal} , which has better quality. We then further fine-tune the model with the new labels using the same objective function as the baseline detection

model, e.g., binary cross-entropy loss [19]. More details on our CAL algorithm are provided in Appendix A.1.

Integrating CAL with existing methods. CAL can be integrated with existing cloud detection models without altering the design or adding any learnable parameters (see Figure 2(b)). CAL can be applied directly to any existing model without retraining. However, retraining with CAL has demonstrated enhanced results.

4 Experiments

Dataset. In our experiments, we utilize the publicly accessible 38-Cloud dataset [20]. This dataset comprises 38 Landsat-8 scenes, each with dimensions of 1000×1000 pixels, and includes manually generated pixel-level cloud mask ground truths, albeit with some noise. The dataset is structured into numerous patches, each measuring 384×384 pixels, and is composed of four spectral channels: Red, Green, Blue, and Near Infrared (NIR). Specifically, there are 8,400 patches designated for training purposes and an additional 9,201 patches allocated for testing.

Implementation details. We trained our CAL model using the Adam optimizer [21] with a learning rate of 0.001 for 3 epochs. The initial threshold value was set to 60.0, with a step size of 2.0. We defined the best loss as infinity and the lower bound as 45.0. Additionally, we updated the threshold every 150 steps. As an illustration, we integrated CAL with the widely used U-Net [22], DeeplabV3 [23], and FCN [24] segmentation models.

Results. Table 1 reveals the following insights: (1) Our baseline models, U-Net, DeeplabV3, and FCN, perform notably worse than specialized methods like LWCDnet. This discrepancy suggests the presence of unique challenges in cloud segmentation. (2) Encouragingly, our CAL approach significantly enhances the performance of U-Net, DeeplabV3, and FCN, demonstrating an improvement of 20% or more in mIoU. This achievement positions our method as the top-performing solution among all competitors, underscoring its effectiveness in mitigating the challenges posed by noisy cloud labeling. In particular, U-Net+CAL achieves the best performance among all the competitors. (3) In addition to this, our proposed CAL consistently improves the performance of FCN, DeeplabV3, and U-Net-based cloud detection models by simply plugging in without any design changes, thus making our algorithm model agnostic in nature.

Table 1: Performance comparison on the 38-Cloud dataset.

Method	mIoU	Precision	Recall	F1-Score	OA
CD-CTFM [5]	84.13	91.09	89.22	90.15	95.45
CD-AttDLV3+ [17]	81.24	88.85	87.58	88.21	94.49
CloudAttU [25]	84.73	90.62	89.95	90.28	95.92
CloudFCN [18]	83.31	88.81	89.61	89.21	95.66
CD-Net [16]	89.70	94.30	94.70	94.50	95.40
LWCDnet [4]	89.90	95.10	94.30	94.70	95.30
FCN [24]	60.40	80.69	63.02	69.07	92.38
DeeplabV3 [23]	44.02	86.25	89.53	86.73	96.07
U-Net [22]	73.69	74.31	97.17	82.10	91.20
FCN + CAL	83.94	91.52	87.73	89.26	96.74
DeeplabV3 + CAL	83.95	85.95	94.17	89.38	96.77
U-Net + CAL	93.15	96.37	96.67	96.44	98.61

5 Conclusion

In this study, we tackle a key challenge associated with noisy training data in cloud analysis, specifically focusing on cloud segmentation in remote sensing. We propose a novel model-agnostic self-labeling approach that iteratively enhances the quality of training data annotations, resulting in improved cloud segmentation model performance. Our method is designed to seamlessly integrate with various existing segmentation approaches. Extensive experiments conducted on a widely recognized cloud segmentation benchmark validate the effectiveness of our approach. Our method outperforms existing methods, establishing a new state-of-the-art performance level.

References

- [1] Zhengqiang Li, Ying Zhang, Jie Shao, Baosheng Li, Jin Hong, Dong Liu, Donghui Li, Peng Wei, Wei Li, Lei Li, Fengxia Zhang, Jie Guo, Qian Deng, Bangxin Wang, Chaolong Cui, Wanchun Zhang, Zhenzhu Wang, Leiku Yang, Hua Xu, Xingfeng Chen, Li Li, and Lili Qie. Remote sensing of atmospheric particulate mass of dry PM_{2.5} near the ground: Method validation using ground-based measurements. *Remote Sensing of Environment*, 173:59–68, 2 2016.
- [2] William B. Rossow and Leonid C. Garder. Cloud detection using satellite measurements of infrared and visible radiances for ISCCP. *Journal of Climate*, 6(12):2341–2369, 12 1993.
- [3] Liyuan Li, Xiaoyan Li, Linyi Jiang, Xin Su, and Fansheng Chen. A review on deep learning techniques for cloud detection methodologies and challenges. *Signal, Image and Video Processing*, 15(7):1527–1535, 4 2021.
- [4] Chen Luo, Shanshan Feng, Xiaofei Yang, Yunming Ye, Xutao Li, Baoquan Zhang, Zhihao Chen, and Yingling Quan. LWCDNet: a lightweight network for efficient cloud detection in remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, 60:1–16, 1 2022.
- [5] Weigong Ge, Xiuhai Yang, and Li Zhang. CD-CTFM: a lightweight CNN-Transformer network for remote sensing cloud detection fusing multiscale features. *arXiv (Cornell University)*, 6 2023.
- [6] Anna Heinle, Andreas Macke, and Anand Srivastav. Automatic cloud classification of whole sky images. *Atmospheric Measurement Techniques*, 3(3):557–567, 5 2010.
- [7] David Frantz, Erik Haß, Andreas Uhl, Johannes Stoffels, and Joachim Hill. Improvement of the Fmask algorithm for Sentinel-2 images: Separating clouds from bright surfaces based on parallax effects. *Remote Sensing of Environment*, 215:471–481, 9 2018.
- [8] Chengquan Huang, Nancy E. Thomas, Samuel N. Goward, Jeffrey G. Masek, Zhiliang Zhu, J. R. G. Townshend, and James E. Vogelmann. Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *International Journal of Remote Sensing*, 31(20):5449–5464, 10 2010.
- [9] Zhe Zhu and Curtis E. Woodcock. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118:83–94, 3 2012.
- [10] Xiatian Zhu, Chen Change Loy, and Shaogang Gong. Video synopsis by heterogeneous multi-source correlation. In *IEEE International Conference on Computer Vision*. IEEE, 2013.
- [11] Xiatian Zhu, Chen Change Loy, and Shaogang Gong. Constructing robust affinity graphs for spectral clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1450–1457, 2014.
- [12] Yanlin Sui, Bin He, and Tiantian Fu. Energy-based cloud detection in multispectral images based on the SVM technique. *International Journal of Remote Sensing*, 40(14):5530–5543, 2 2019.
- [13] Jing Wei, Wei Huang, Zhanqing Li, Lin Sun, Xiaolin Zhu, Qiangqiang Yuan, Lei Liu, and Maureen Cribb. Cloud detection for Landsat imagery by combining the random forest and superpixels extracted via energy-driven sampling segmentation approaches. *Remote Sensing of Environment*, 248:112005, 10 2020.
- [14] Douglas Chai, Shawn Newsam, Hankui K. Zhang, Yifan Qiu, and Jingfeng Huang. Cloud and cloud shadow detection in Landsat imagery based on deep convolutional neural networks. *Remote Sensing of Environment*, 225:307–316, 5 2019.
- [15] Fengying Xie, Mengyun Shi, Zhenwei Shi, Jingwei Yin, and Danpei Zhao. Multilevel cloud detection in remote sensing images based on deep learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(8):3631–3640, 8 2017.
- [16] Jingyu Yang, Jianhua Guo, Huanjing Yue, Zhiheng Liu, Haofeng Hu, and Kun Li. CDNET: CNN-Based Cloud Detection for Remote sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 57(8):6195–6211, 8 2019.
- [17] Xudong Yao, Qing Guo, and An Liu. Light-Weight Cloud Detection Network for Optical Remote Sensing Images with Attention-Based DeeplabV3+ Architecture. *Remote Sensing*, 13(18):3617, 9 2021.

- [18] Alistair Francis, Panagiotis Sidiropoulos, and Jan-Peter Muller. CloudFCN: Accurate and Robust Cloud Detection for Satellite Imagery with Deep Learning. *Remote Sensing*, 11(19):2312, 10 2019.
- [19] Zhilu Zhang and Mert R. Sabuncu. Generalized cross entropy loss for training deep neural networks with noisy labels. *arXiv (Cornell University)*, 31:8792–8802, 12 2018.
- [20] Sorour Mohajerani, Thomas Krammer, and Parvaneh Saeedi. A Cloud Detection Algorithm for Remote Sensing Images Using Fully Convolutional Neural Networks. *2018 IEEE 20th International Workshop on Multimedia Signal Processing (MMSP)*, pages 1–5, 8 2022.
- [21] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv (Cornell University)*, 12 2014.
- [22] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-NET: Convolutional Networks for Biomedical Image Segmentation*. 1 2015.
- [23] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking Atrous convolution for semantic image segmentation. *arXiv (Cornell University)*, 6 2017.
- [24] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):640–651, 4 2017.
- [25] Yanan Guo, Xiaoqun Cao, Bainian Liu, and Mei Gao. Cloud detection for satellite imagery using Attention-Based U-Net convolutional neural network. *Symmetry*, 12(6):1056, 6 2020.

A Appendix

A.1 CAL: Algorithm details

We illustrate the dynamic thresholding-based CAL algorithm in detail. Here p_{th} refers to the learnable_threshold parameter, M_{cal} refers to the new_labels parameter, I refers to the raw input satellite image images respectively. For each training iteration over a mini-batch, we optimize the learnable_threshold parameter based on the loss criterion performance. If the loss is higher, this indicates that the threshold is overpowering else it is underpowering. Based on this, our CAL algorithm reduces/increases the threshold by a fixed incremental threshold δ_x respectively.

Algorithm 1: Cloud Adaptive Labeling

```
1 learnable_threshold = 60.0 // initial threshold
2 delta_x = 2.0 // step size
3 best_loss = float('inf')
4 lower_bound = 45.0
5 update_frequency = 150
6 while not done do
7   new_labels = binarize_image(images, threshold=learnable_threshold)
8   outputs = model(images)
9   loss = criterion(outputs, new_labels)
10  best_loss = min(best_loss, loss)
11  if (idx + 1) % update_frequency == 0 then
12    if loss > best_loss then
13      learnable_threshold -= delta_x // Decrease the threshold
14      learnable_threshold = max(lower_bound, learnable_threshold)
15    else
16      learnable_threshold += delta_x // Increase the threshold
```

Model agnosticity of CAL. Our proposed CAL works as a plug-and-play module on existing cloud detection models like FCN, Deeplabv3, and U-Net respectively. From the results in Table 2, it is seen that our CAL when plugged in improves consistently over the baseline model without CAL. This is observed for all the models thus proving the model-agnosticity of our algorithm design.

Table 2: Performance comparisons of models with and without CAL.

Methods	mIoU	Precision	Recall	F1-Score	OA
FCN	60.40	80.69	63.02	69.07	92.38
FCN + CAL (ours)	83.94	91.52	87.73	89.26	96.74
U-Net	73.69	74.31	97.17	82.10	91.20
U-Net + CAL (ours)	93.15	96.37	96.67	96.44	98.61
DeeplabV3	44.02	86.25	89.53	86.72	96.07
DeeplabV3 + CAL (ours)	83.95	85.95	94.18	89.39	96.77

Training convergence due to CAL. Existing labels provided by the model often have incorrect cloud segmentation annotation as shown in Fig 1. Such faulty annotations make the model slower to converge. As seen from Fig 3, the variant with CAL (Fig 3 (b)) converges faster than the variant without CAL (Fig 3 (a)). This shows the efficacy of our adaptive thresholding improving the labels. Additionally, our CAL improves the stability and accuracy of the model precision than the baseline variant as shown from Fig 3 (c,d) respectively.

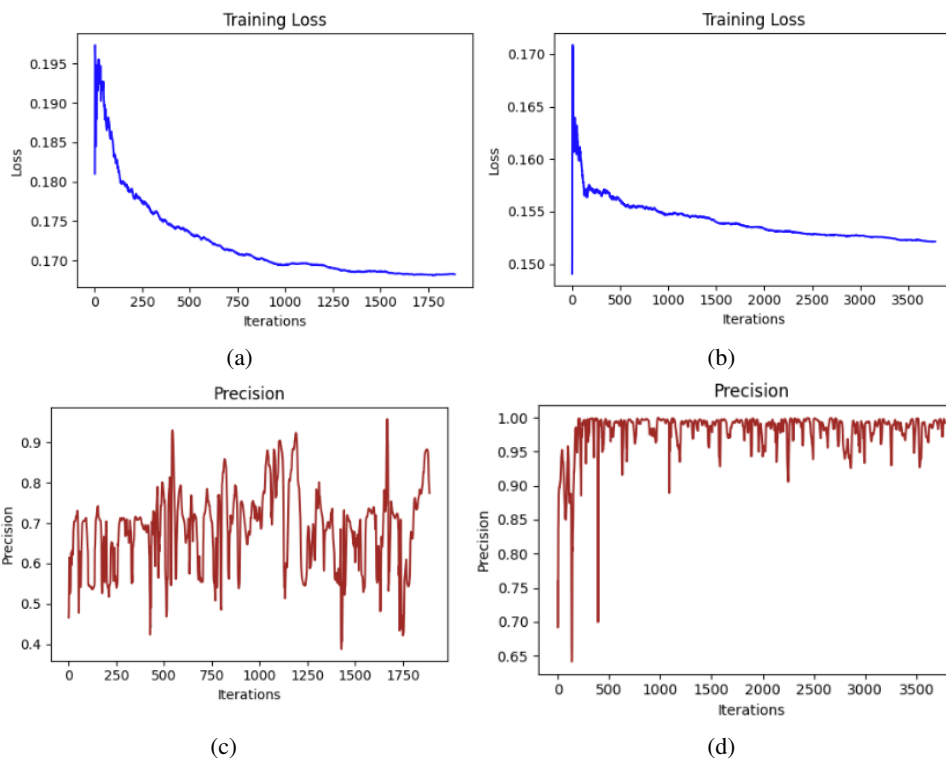


Figure 3: (a) Training loss curve for finetuning U-Net (b) Training loss curve for fine-tuning U-Net with CAL. Finetuning with CAL results in faster convergence than fine-tuning without CAL (c) Precision curve for finetuning U-Net (d) Precision curve for fine-tuning U-Net with CAL. Finetuning with CAL quickly stabilizes the precision as opposed to fine-tuning without CAL.