Multi-label Cloud Segmentation
Using a Deep Network

Soumyabrata Dev[‡], Shilpa Manandhar[*], Yee Hui Lee[*], and Stefan Winkler[§]

[‡] ADAPT SFI Research Centre, Trinity College Dublin, Ireland
[*] School of Electrical and Electronic Engineering, Nanyang Technological University (NTU), Singapore
[§] School of Computing, National University of Singapore (NUS)

Abstract—**Different empirical models have been developed for cloud detection. There is a growing interest in using the ground-based sky/cloud images for this purpose. Several methods exist that perform binary segmentation of clouds. In this paper, we propose to use a deep learning architecture (U-Net) to perform multi-label sky/cloud image segmentation. The proposed approach outperforms recent literature by a large margin.**

I. INTRODUCTION

Clouds play an important role in air-to-ground communication. In particuler, satellite communications via very small aperture terminal operating in the Ka-band (20/30 GHz) can be severely impaired by clouds. Cloud attenuation increases as the operating frequency increases. In the literature, there are several works that discuss efficient methods to detect clouds either using empirical models [1] or via satellite images and ground-based camera images. However, with the recent advancement ground-based sensing and photogrammetric techniques, ground-based observations using sky cameras are gaining more popularity, owing to the lower operating cost as well as higher temporal and spatial image resolution. Several methods have been devised and proposed in the literature that attempt to perform a binary detection of labels – *sky* and *cloud* in ground-based sky camera images. These methods [2]–[4] typically use a discriminatory color feature in different color spaces and components for the cloud detection.

Recently, Dev et al. [5] proposed using multi-labels for whole sky images by classifying them further into *sky*, *thin cloud*, and *thick cloud*, by modelling the images as a continuous-valued multi-variate distribution. This approach provided good detection rates for *sky* and *thick cloud*, but *thin cloud* detection rate was poor.

II. MULTI-LABEL SEGMENTATION OF CLOUD IMAGES

Currently, with the availability of more powerful computing resources, deep neural networks are used for various tasks remote sensing. Deep convolutional neural nets have shown impressive results for the task of image segmentation. We employ a popular deep learning architecture called U-Net [6] for the task of recognizing the different labels in a sky/cloud image.[1]

U-Net is a deep convolutional neural network that was originally designed for accurate segmentation of electron microscopy images. We train the U-Net architecture on 128×128 resized images from our dataset. The output of the U-Net model is a probabilistic mask, wherein each pixel has a *softargmax* value in the range $[0, 1]$. This value indicates the degree of cloudiness of the pixel in the sky/cloud image. Finally, we threshold this probabilistic mask into a ternary map using thresholds of 0.3 and 0.6. We use these equally spaced thresholds considering equal misclassification costs to the individual labels.

III. EXPERIMENTS & RESULTS

A. Dataset

We use the dataset introduced in our previous work [5]. It is based on the images of the HYTA dataset [2], from which we generated 32 cropped sky/cloud images, ensuring that diverse sky scene types are being represented. In consultation with expert from the Meteorological Service Singapore, we also annotated the images with segmentation ground-truth containing multiple labels – *sky*, *thin cloud* and *thick cloud*.[2]
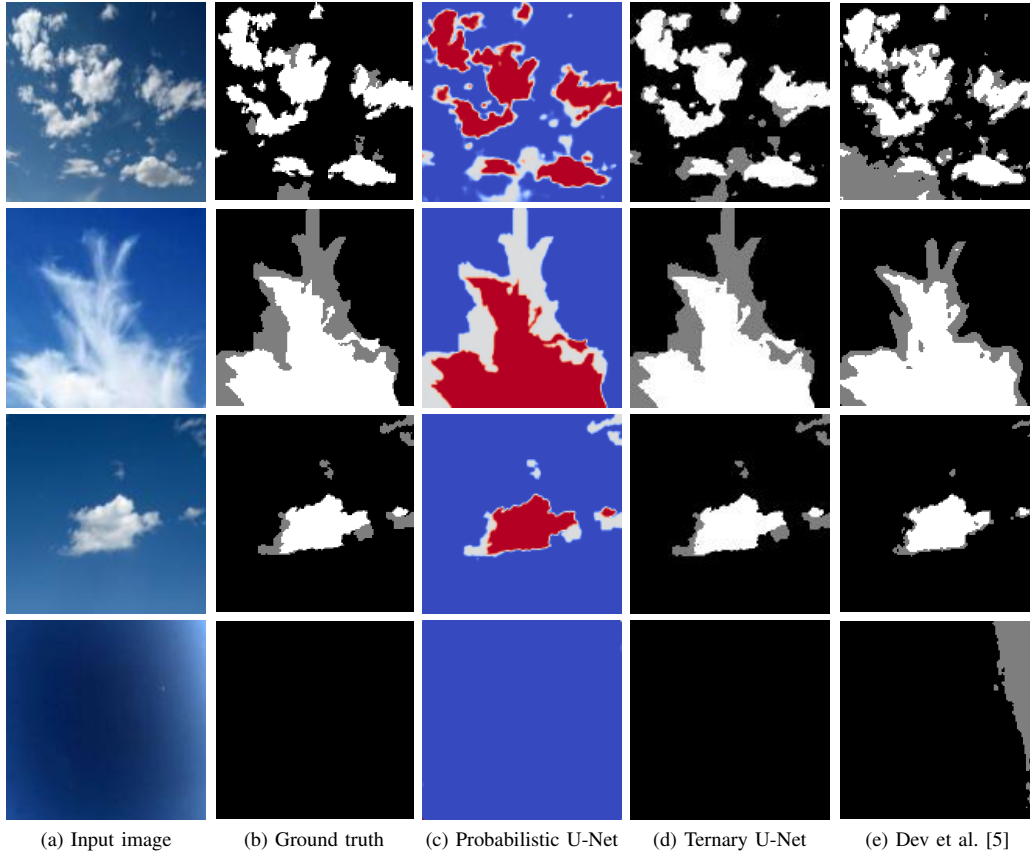
B. Qualitative Evaluation

Figure 1 provides a qualitative evaluation of our proposed method using the U-Net architecture. We benchmark our approach against the previous approach [5]. We observe that our approach can efficiently detect the various labels in the sky/cloud image. Most of the improvements are observed in the *thin cloud* regions, which are notoriously difficult to detect using traditional approaches. Also, our deep-learning based approach significantly reduces false positives in the event of a clear sky scene. Furthermore, U-Net also produces a probabilistic output, that can assist the remote sensing analysts to incorporate their subjective judgement in the generated results (although the variation of pixel values within a single cluster is quite small in these examples).

C. Quantitative Evaluation

In this section, we provide a quantitative evaluation of our proposed approach. We perform a random split of training and testing sets in the ratio of $80 : 20$. We do not employ any image augmentation techniques during training. The U-Net model is

---

Send correspondence to Y. H. Lee, E-mail: EYHLee@ntu.edu.sg.

[1]The code of all simulations in this paper is available online at https://github.com/Soumyabrata/multilabel-unet

[2]More details can be found here: https://github.com/Soumyabrata/HYTA

| (a) Input image | (b) Ground truth | (c) Probabilistic U-Net | (d) Ternary U-Net | (e) Dev et al. [5] |

Fig. 1: Qualitative evaluation of our proposed approach w.r.t. other existing approaches. We use a diverging color map of *coolwarm* for visualizing the probabilistic output of U-Net – the degree of *redness* indicate higher probability of cloudiness.

TABLE I: Performance evaluation of the proposed approach on the dataset. Values show the misclassification error w.r.t. the three labels. We report the error percentage (%) in the detection of sky-, thin-cloud- and thick-cloud- labels of the testing images.

Approach	Sky	Thin cloud	Thick cloud
Dev et al. [5]	15.4%	52.0%	23.4%
Proposed approach	7.3%	4.4%	4.4%

trained on the training set and evaluated on the testing set. We perform this experiment 10 times, in order to remove any sampling bias. We employ the same experimental procedure to evaluate the Dev et al. [5] approach.

We compute the error percentages of the individual labels – *sky*, *thin cloud*, and *thick cloud* – across all the testing images. Subsequently, we compute the average error percentage of the different labels over the set of 10 experiments. Table I summarizes the results. Despite the relatively small size of the dataset, we observe that the performance of our proposed approach is significantly better than the previous work for all the three labels. By far the biggest improvement is observed in the detection of *thin cloud*.

IV. CONCLUSION

In this paper, a deep learning architecture is used for multi-label cloud segmentation. Our proposed approach represents a significant improvement over previous work, especially for thin cloud detection. This is particularly interesting for remote sensing applications [7], as thin clouds are often mislabelled using conventional approaches.

REFERENCES

[1] F. Yuan, Y. H. Lee, Y. S. Meng, and J. T. Ong, "Water vapor pressure model for cloud vertical structure detection in tropical region," *IEEE Trans. Geosci. Rem. Sens.*, vol. 54, no. 10, pp. 5875–5883, Oct 2016.

[2] Q. Li, W. Lu, and J. Yang, "A hybrid thresholding algorithm for cloud detection on ground-based color images," *Journal of Atmospheric and Oceanic Technology*, vol. 28, no. 10, pp. 1286–1296, Oct. 2011.

[3] M. P. Souza-Echer, E. B. Pereira, L. S. Bins, and M. A. R. Andrade, "A simple method for the assessment of the cloud cover state in high-latitude regions by a ground-based digital camera," *Journal of Atmospheric and Oceanic Technology*, vol. 23, no. 3, pp. 437–447, Mar. 2006.

[4] S. Dev, Y. H. Lee, and S. Winkler, "Systematic study of color spaces and components for the segmentation of sky/cloud images," in *Proc. International Conference on Image Processing (ICIP)*, 2014, pp. 5102–5106.

[5] S. Dev, Y. H. Lee, and S. Winkler, "Multi-level semantic labeling of sky/cloud images," in *Proc. International Conference on Image Processing (ICIP)*, 2015, pp. 636–640.

[6] O. Ronneberger, P. Fischer, and T. 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.

[7] S. Manandhar, S. Dev, Y. H. Lee, and S. Winkler, "Predicting GPS-based PWV measurements using exponential smoothing," in *Proc. IEEE AP-S Symposium on Antennas and Propagation and USNC-URSI Radio Science Meeting*, 2019.