TRULY GENERATIVE DATA AUGMENTATION FOR IMAGE SEGMENTATION - CASE OF CLOUD IMAGES

Anonymous authors
Paper under double-blind review

ABSTRACT

Deep learning based approaches have been instrumental in dealing with the problem of supervised image segmentation. However, they require a large amount of annotated data. Although previous approaches either use the limiting label-invariant transformations or are not reliable for generative augmentation, this paper proposes an organic approach to generatively perform data augmentation for image segmentation task in the case of sky/cloud images. It works by separately generating the background clear sky image and the foreground cloud masks from two separately trained GAN architectures, which are later merged together.

1 INTRODUCTION

Recent years have seen great success with deep learning-based approaches for supervised learning tasks (Hatamizadeh et al., 2022; Ye et al., 2022). However, this comes with an expensive prerequisite of largely annotated and balanced datasets. Consequently, various data augmentation strategies have been explored in the literature with great success (Antoniou et al., 2017). A typical data augmentation approach uses existing labeled data points to generate new ones. This is conditioned on the constraints that the ground truth annotations of the augmented data must be reliable and that the augmented data points fall in a similar distribution to that of the original.

Data augmentation approaches for generating new data points can be classified as either label-invariant transformations of the original image (Ratner et al., 2017); or to generate entirely new images and their corresponding ground truth annotations. Although the former approach adheres to the augmentation constraints, it is limited by the cardinality of the original dataset (Cubuk et al., 2019). On the other hand, purely generative approaches can negate this limitation but have a greater difficulty meeting the constraints. Two types of generative augmentation approaches have been discussed in the literature. Bowles et al. (2018) appends the ground truth segmentation mask as an extra channel to the original image before training the generative network. Although this approach was found to be experimentally useful, it does not ensure the reliability of the generated annotated mask with respect to the generated image. Another approach is to generate/estimate a segmentation mask. Sandfort et al. (2019) uses CycleGAN Zhu et al. (2017), a style transfer model, for the job, while (Jain et al., 2021) uses K-Means, an unsupervised clustering approach, to estimate segmentation masks for the DCGAN Radford et al. (2015) generated images. However, this approach leads to a chicken-and-egg problem. If the method used to estimate the segmentation mask is reliable, then there is no need to increase the data and train a supervised model for image segmentation in the first place. And if the method is not reliable, then the corresponding segmentation mask must not be used as a ground truth label in the augmented dataset.

Consequently, to navigate the aforementioned problems in prevalent generative approaches, this paper proposes a novel and organic way to deal with the problem in the case of sky/cloud image segmentation. This is done by training two separate GANs for cloud masks and background sky, thereby ensuring that the segmentation maps can be used as ground truth.

2 CASE OF CLOUD IMAGE SEGMENTATION

Images from ground-based sky camera systems have been widely used to analyze cloud movements with high temporal and spatial resolution (Siddiqui et al., 2019). However, due to the lack of proper

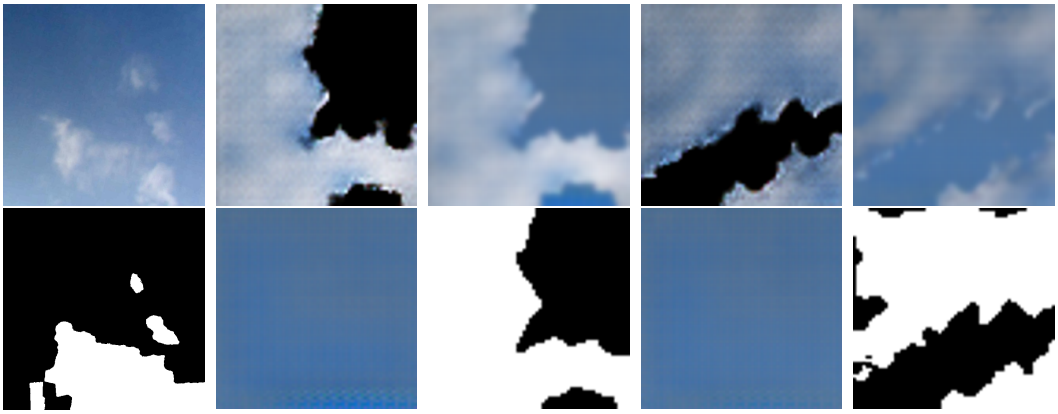
structure and fuzzy boundaries, it is not always easy to accurately segment these images. SWIMSEG (Dev et al., 2016), containing 1013 sky/cloud images (say, $\mathbb{C}$) with manually annotated binary segmentation maps (say, $\mathbb{G}$), is one of the most widely used dataset for the task (Pierce et al., 2022). Here, the task is to propose a truly generative data augmentation strategy for this dataset.

2.1 Proposed Methodology

The paper formulizes the problem as a human artist who is tasked to perform the augmentation task. Organically, the first step would be to paint a clear sky and then to paint clouds over it. However, while painting the clouds, the artist can place a tracing paper underneath to get an impression of the annotation mask. Accordingly, given an image $\boldsymbol{I} \in \mathbb{C}$ and the corresponding ground truth segmentation map $\boldsymbol{J} \in \mathbb{G}$, all pixels that are 0 in $\boldsymbol{J}$ were also made 0 in all three channels of $\boldsymbol{I}$. Now, a DCGAN is trained on the resulting set $\widetilde{\mathbb{C}}$ for 5000 epochs. Let the image generated by the generator of this model be represented by $\boldsymbol{M}$.

To draw the clear sky background, another DCGAN (similar to the previous one) is trained on the 'sky' class of the SWIMCAT (Dev et al., 2015) dataset. SWIMCAT has images similar to those of SWIMSEG, except that in this dataset the sky/cloud images are classified based on the types of cloud in the image. Let the image generated by the generator of this model be represented by $\boldsymbol{N}$.

Now, if a pixel value is less than the Otsu threshold (Otsu, 1979) in the three channels of $\boldsymbol{M}$, they are substituted by the pixel value of $\boldsymbol{N}$. At the same time, for the ground truth annotation image, say $\boldsymbol{P}$, the pixel value is set to 0, otherwise it is set to 1. Finally, both images are slightly smoothed to get the desired pair ready for image augmentation (as shown in Figure 1 (c,e)).



(a) SWIMSEG (b) Sample result 1 (c) Sample result 1 (d) Sample result 2 (e) Sample result 2

Figure 1: (a) Shows a sample image and its ground truth map from the SWIMSEG dataset; (b,d) show raw DCGAN generated images for cloud mask and clear sky background; (c,e) show the finally generated sky/cloud image and the corresponding annotation binary segmentation map.

3 Conclusion and Future Work

Prevalent approaches for data augmentation in image segmentation task are limited by the cardinality of the originally annotated set, or are not generatively reliable with respect to the annotation. On the contrary, this paper proposes a truly generative approach using the standard DCGAN models. One DCGAN is responsible for generating only the cloud mask, while the other one is responsible for generating only the background sky/cloud image. These two generated images are then combined to generate sky/cloud images and their corresponding binary segmentation maps. The annotations generated as such are bound to match the generated cloud profiles. While this paper presents a working proof for the case of sky/cloud images, the idea needs to be further improved for a broader applicability towards data augmentation for image segmentation problems. Moving forward, some context exchange between the two DCGANs might help in conherent generation of blue shades in the cloud masks and the background clear sky images. Additionally, post-processing can be improved for a seamless integration of the two.

URM STATEMENT

The authors acknowledge that at least one key author of this work meets the URM criteria of the ICLR 2023 Tiny Papers Track.

CODE RELEASE

For reproducibility, the codebase used for the experiments conducted in this paper will be released at `https://github.com/XXXX`. The link is currently redacted to maintain anonymousity of the authors for double blind peer review.

REFERENCES

Antreas Antoniou, Amos Storkey, and Harrison Edwards. Data augmentation generative adversarial networks. *arXiv preprint arXiv:1711.04340*, 2017.

Christopher Bowles, Liang Chen, Ricardo Guerrero, Paul Bentley, Roger Gunn, Alexander Hammers, David Alexander Dickie, Maria Valdés Hernández, Joanna Wardlaw, and Daniel Rueckert. Gan augmentation: Augmenting training data using generative adversarial networks. *arXiv preprint arXiv:1810.10863*, 2018.

Ekin D Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc V Le. Autoaugment: Learning augmentation strategies from data. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 113–123, 2019.

Soumyabrata Dev, Yee Hui Lee, and Stefan Winkler. Categorization of cloud image patches using an improved texton-based approach. In *2015 IEEE International Conference on Image Processing (ICIP)*, pp. 422–426, 2015. doi: 10.1109/ICIP.2015.7350833.

Soumyabrata Dev, Yee Hui Lee, and Stefan Winkler. Color-based segmentation of sky/cloud images from ground-based cameras. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(1):231–242, 2016.

Ali Hatamizadeh, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Bennett Landman, Holger R Roth, and Daguang Xu. UNETR: Transformers for 3D medical image segmentation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 574–584, 2022.

Mayank Jain, Conor Meegan, and Soumyabrata Dev. Using GANs to augment data for cloud image segmentation task. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 3452–3455, 2021. doi: 10.1109/IGARSS47720.2021.9554993.

Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66, 1979.

Benjamin G. Pierce, Joshua S. Stein, Jennifer L. Braid, and Daniel Riley. Cloud segmentation and motion tracking in sky images. *IEEE Journal of Photovoltaics*, 12(6):1354–1360, 2022. doi: 10.1109/JPHOTOV.2022.3215890.

Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.

Alexander J Ratner, Henry Ehrenberg, Zeshan Hussain, Jared Dunnmon, and Christopher Ré. Learning to compose domain-specific transformations for data augmentation. *Advances in neural information processing systems*, 30, 2017.

Veit Sandfort, Ke Yan, Perry J. Pickhardt, and Ronald M. Summers. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Scientific Reports*, 9(1):16884, Nov 2019. ISSN 2045-2322. doi: 10.1038/s41598-019-52737-x. URL `https://doi.org/10.1038/s41598-019-52737-x`.

Talha Ahmad Siddiqui, Samarth Bharadwaj, and Shivkumar Kalyanaraman. A deep learning approach to solar-irradiance forecasting in sky-videos. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 2166–2174, 2019. doi: 10.1109/WACV.2019.00234.

Liang Ye, Zhiguo Cao, Zhibiao Yang, and Huasong Min. CCAD-Net: A cascade cloud attribute discrimination network for cloud genera segmentation in whole-sky images. *IEEE Geoscience and Remote Sensing Letters*, 19:1–5, 2022.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.