Accuracy Analysis of using Intersection over Union at Normalized Difference Vegetation Index

Hyun Choi (1)

[1] Department of Civil Engineering, Kyungnam Univ., 7 Gyeongnamdaehak-ro, Masanhappo-gu, Changwon-si, Gyeongsangnam-do 51767, Korea

Email: hchoi@kyungnam.ac.kr

Abstract: In South Korea, it is very important to build accurate databases of forests, which occupy most of the territory, in order to prevent disasters and accidents. This paper is about that comparative analysis of generalized intersection over union and error matrix in NDVI. The Normalized Difference Vegetation Index (NDVI) depends largely on the image classification and accuracy setting method. In recent remote exploration, image classification is being carried out based on deep learning. Currently, almost all remote sensing uses error matrix for accuracy analysis, but in the field of image processing in computer science, IoU is used for accuracy analysis. Therefore, the accuracy is analyzed using the IoU method, but the difference in accuracy levels between the Vegetation Index identified using the NDVI and that identified using the deep learning technique cannot be known. Intersection over Union (IoU) is the most popular evaluation metric used in the object detection benchmarks. IoU, also known as Jaccard index, is the most commonly used metric for comparing the similarity between two arbitrary shapes. Therefore, to solve this problem, in this study, the Vegetation Index was calculated using the deep learning method and the accuracy of the Vegetation Index calculated using the existing error matrix and that using the IoU method were compared. To extract objects from satellite images, conifers and deciduous trees were identified using the semantic image classification technique in the CNN(Convolutional Neural Networks) technique using a DeepLab model. DeepLab models are based on atrous convolution. These models have the advantage of easy image classification since their receptive field is wide and carry out calculations as if unpooling and convolution were combined. To evaluate the performance of the model, the accuracy of classification was calculated using IoU, and the results were shown.

Keywords: Normalized Difference Vegetation Index(NDVI), DeepLab model, CNN (Convolutional Neural Networks), Intersection over Union (IoU)

1. Intorduction

To determine the accuracy of image classification, it is checked using the error matrix technique. How ever, when the accuracy is analyzed with the error matrix, which is based on experience points, the reli ability of the accuracy declines. Therefore, the accuracy is analyzed using the IoU method, but the diff erence in accuracy levels between the Vegetation Index identified using the NDVI and that identified using the deep learning technique cannot be known. Intersection over Union (IoU) is the most IoU, als o known as Jaccard index, is the most commonly used metric for comparing the similarity between tw o arbitrary shapes. Therefore, to solve this problem, in this study, the Vegetation Index was calculated using the deep learning method and the accuracy of the Vegetation Index calculated using the existing error matrix and that using the IoU method were compared.

2. Generalized Intersection over Union

Intersection over Union (IoU) for comparing similarity between two arbitrary shapes (volumes) A,B $\subseteq \mathbb{S} \in \mathbb{R}^{\boldsymbol{n}}$ is attained by :

$$\mathbf{IoU} = \frac{|\boldsymbol{A} \cap \boldsymbol{B}|}{|\boldsymbol{A} \cup \boldsymbol{B}|} \tag{6}$$

Two appealing features, which make this similarity measure popular for evaluating many 2D/3D computer vision tasks are as follows:

Algorithm 1 : Generalized Intersection over Union

Input : Two arbitrary convex shapes : A,B $\subseteq \mathbb{S} \in \mathbb{R}^{\boldsymbol{n}}$

Output : GIoU

For A and B, find the smallest enclosing convex object C, where C $\subseteq \mathbb{S} \in \mathbb{R}^{\boldsymbol{n}}$

$$\mathbf{IoU} = \frac{|\boldsymbol{A} \cap \boldsymbol{B}|}{|\boldsymbol{A} \cup \boldsymbol{B}|}$$

$$\text{GIoU} = \text{IoU} - \frac{|\boldsymbol{C}/(\boldsymbol{A} \cup \boldsymbol{B})|}{|\boldsymbol{C}|}$$

IoU as a distance, $\mathbf{e.g.}\, \boldsymbol{\mathcal{L}_{IoU}} = \mathbf{1} - \boldsymbol{IoU}$, is a metric (by mathematical definition) [1]. It means $\boldsymbol{\mathcal{L}_{IoU}}$ fulfills all properties of a metric such as non-negativity, identity of indiscernibles, symmetry and triangle inequality.
IoU is invariant to the scale of the problem. This means that the similarity between two arbitrary shapes A and B is independent from the scale of their space S (the proof is provided in supp. material).

However, IoU has a major weakness :
If $|\boldsymbol{A} \cap \boldsymbol{B}|$ = 0, IoU(A, B) = 0. In this case, IoU does not reflect if two shapes are in vicinity of each other or very far from each other.

To address this issue, we propose a general extension to IoU, namely Generalized Intersection over Union GIoU. For two arbitrary convex shapes (volumes) A,B $\subseteq \mathbb{S} \in \mathbb{R}^{\boldsymbol{n}}$, we first find the smallest convex shapes C $\subseteq \mathbb{S} \in \mathbb{R}^{\boldsymbol{n}}$ enclosing both A and B. For comparing two specific types of geometric shapes, C can be from the same type. For example, two arbitrary ellipsoids, C could be the smallest ellipsoids enclosing them. Then we calculate a ratio between the volume (area) occupied by C excluding A and B and divide by the total volume (area) occupied by C. This represents a normalized measure that focuses on the empty volume (area) between A and B. Finally GIoU is attained by subtracting this ratio from the IoU value. The calculation of GIoU is summarized in Alg.
GIoU as a new metric has the following properties :

Similar to IoU, GIoU as a distance, $\mathbf{e.g.}\, \boldsymbol{\mathcal{L}_{IoU}} = \mathbf{1} - \boldsymbol{IoU}$, holding all properties of a metric such as non-negativity, identity of indiscernibles, symmetry and triangle inequality.

Similar to IoU, GIoU is invariant to the scale of the problem.
GIoU is always a lower bound for IoU, i.e. $\forall A, B \subseteq \mathbb{S}$ GIoU(A,B) ≤ IoU(A,B), and this lower bound becomes tighter when A and B have a stronger shape similarity and proximity, i.e. $\lim_{A \to B} \boldsymbol{GIoU(A,B) = IoU(A,B)}$.
$\forall A, B \subseteq \mathbb{S}$, 0 ≤ IoU(A,B) ≤ 1, but GIoU has a symmetric range, i.e. $\forall A, B \subseteq \mathbb{S}$, $-1 \leq$ GIoU(A,B) ≤ 1.

Similar to IoU, the value 1 occurs only when two objects overlay perfectly, i.e. if $|\boldsymbol{A} \cup \boldsymbol{B}| = |\boldsymbol{A} \cap \boldsymbol{B}|$, then GIoU = IoU = 1

GIoU value asymptotically converges to -1 when the ratio between occupying regions of two shapes, $|\boldsymbol{A} \cup \boldsymbol{B}|$, and the volume (area) of the enclosing shape $|\boldsymbol{C}|$ tends to zero, i.e. $\lim_{\frac{|A \cup B|}{|C|} \to 0} \boldsymbol{GIoU(A,B) = -1}$

In summary, this generalization keeps the major properties of IoU while rectifying its weakness. Therefore, GIoU can be a proper substitute for IoU in all performance measures used in 2D/3D computer vision tasks. In this paper, we only focus on 2D object detection where we can easily derive an analytical solution for GIoU to apply it as both metric and loss. The extension to non-axis aligned 3D cases is left as future work.

3. Accuracy analysis

To extract objects from satellite images, conifers and deciduous trees were identified using the semantic image classification technique in the CNN technique using a DeepLab model. DeepLab models are based on atrous convolution. These models have the advantage of easy image classification since their receptive field is wide and carry out calculations as if unpooling and convolution were combined.

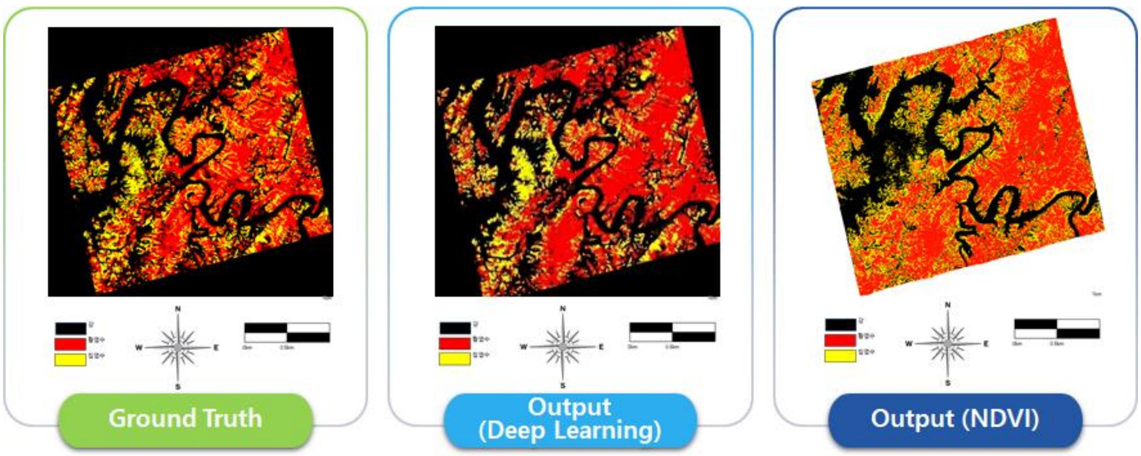

Figure 1. The results of classification of trees into coniferous and deciduous trees for the test data

To evaluate the performance of the model, the accuracy of classification was calculated using IoU, and the results were shown. To analyze performance, the results of classification were compared with the conifers and deciduous trees in the land cover map provided by the Environmental Geographic Information Service (EGIS) of the Republic of Korea. According to the results of accuracy analysis using the error matrix, which was based on random points, the accuracy of the NDVI was 82.4% and that of deep learning was 93.7%, with a difference of 11.3%. According to the results of accuracy analysis using IoU, which is based on grids, the accuracy of the NDVI was 49.4% and that of deep learning was 71.6%, with a difference of about 22.2%. Cases of misclassification occur because the pixels of conifers and deciduous trees are similar.

4. Conclusions

According to the results of accuracy analysis using the error matrix, which is based on random points, the accuracy of the NDVI was 82.4% and that of deep learning was 93.7%, with a difference of 11.3 %. According to the results of accuracy analysis using IoU, which is based on grids, the accuracy of th e NDVI was 49.4% and that of deep learning was 71.6%, with a difference of about 22.2%. Therefore, according to the results of the study on deep learning, the accuracy according to IoU was shown to be about 70% in the classification of the same image and about 60% when learned images were applied t o other images. In the analysis of the Vegetation Index, the accuracy of the results of analysis using th e error matrix was shown to be higher according to the accuracy setting method, and it could be seen t hat if the deep learning technique was used, the accuracy was high. In the case of the deep learning tec hnique, studies to improve performance through learning of the regional characteristics of conifers and deciduous trees are considered necessary.

Acknowledgments

This research was supported by a grant (19SIUE-B148326-02) from the KAS 500-1/2 Image Acquisition and Utilization Technology Development Program, funded by the Ministry of Land, Infrastructure and Transport of the Korean government.

References

1. S. Kosub. A note on the triangle inequality for the jaccard distance. arXiv preprint arXiv:161 2.02696, 2016.