Sugarcane classification optimization method based on high resolution satellite remote sensing image of Lovász hinge

Yongning He[1]，Ming Zhu[1]，Yuqing He[1]，Bo Wu[1]

1. Geographic Information Center of Guangxi, Nanning, 530023 China

KEY WORDS: Lovász hinge,Sugarcane classification, Remote sensing image classification, Loss function

ABSTRACT:

In recent years, great progress has been made in the study of semantic segmentation in the field of computer vision. The accuracy of semantic segmentation has been constantly improved, and it has been widely applied in the fields of automatic driving, medical treatment and remote sensing image classification.Semantic segmentation in all kinds of neural network structure has been optimized, according to different segmentation task put forward different loss function and different optimization algorithm to improve the accuracy of classification, such as used in the classification task more softmax cross entropy loss in the sigmoid function is used in the classification task, two different loss functions have a different impact on classification results, at the same time, the training data set imbalance can also cause the precision of classification result deviation.In the task of remote sensing image classification, it is often necessary to extract and classify a variety of different land types, such as road, water system, vegetation, etc., from an image, but sometimes it is also necessary to extract one of the land types.Due to remote sensing image contains abundant spectral information, so the remote sensing image classification task is different from ordinary classification task scenarios, common softmax and sigmoid function, the number can not meet the existing remote sensing image classification task, this requires a combination of specific classification task to adjust and optimize the loss function, to adapt to the different classification task.As a major sugarcane planting province in China, guangxi plays an important role in the development of China's sugar industry. Therefore, it is of great significance to propose sugarcane planting area through high-resolution satellite remote sensing image.But because of guangxi planting condition is complicated and changeable weather condition, often appear cloudy, so in high resolution satellite remote sensing image acquisition and there is still a big challenge on extraction and classification, and on the high rate of satellite remote sensing image texture feature of sugarcane and cassava, corn and other crops of texture feature are similar, therefore in the process of classification will easy to misjudge corn, cassava as sugar cane, which led to a decline in classification accuracy.This paper combined with the extraction of sugarcane planting area based on high-resolution satellite remote sensing images by Jaccard loss of Lovasz hinge, and compared the effects of different loss functions on the accuracy of the results through experiments. Finally, it was concluded that combining Jaccard loss of Lovasz hinge could effectively reduce losses and improve the extraction accuracy of sugarcane planting area.

1. Introduction

In recent years, with the improvement of the level of computer hardware, the theoretical algorithms that could not be realized or were difficult to be realized before have been fully verified, and the deep neural convolution network as the main research method has been widely popularized in computer vision.More and more computer vision algorithms have been gradually applied to remote sensing image classification. By using manually annotated sample data for model training, the classification accuracy has been greatly improved compared with traditional remote sensing image classification methods.

In artificial intelligence and deep learning algorithms, before has not been widely used in the traditional remote sensing image interpretation, need artificial class feature extraction of remote sensing image surface coverage, then through ENVI or interpretation of ERDAS are extracted, and the classification of these software extraction method is based on support vector machine (SVN), random forests, such as classic algorithms, however, limited by artificial subjective judgment, different classes often show different characteristics, under the influence of illumination, resolution, image quality is bigger, excessive reliance on artificial feature extraction(ZHANG, WU, 2017), which can lead to change a new data to lower precision.

---

first author**:** Yongning He（E-mail: hyn886@qq.com）

Because CNN can automatically learn the characteristics of nonlinear and generate high complexity, break through the limitations of the artificial design (WU, CHEN, 2018), because CNN has the characteristic, makes the depth of the convolution network has been widely used in remote sensing image classification, including remote sensing image target detection, change detection, etc., depth of the task to be able to use the remote sensing image convolution network is resolved.In the CNN network plays an important role in the optimization algorithm, good optimization algorithm determines the model can be trained to factor, loss function plays an important role in the optimization algorithm, by adjusting the loss function, makes the result more inclined to sugar cane classification, thus reducing the leak point, enhance the role of the practical production, furthermore, adjust the loss function can also adjust due to the effects of sample distribution is not balanced, so that the model results tuning a more balanced, as far as possible to reduce the effects of uneven data.Therefore, this paper proposed to use the loss function of Lovasz hinge to adjust, minimize the loss of sugarcane classification, adjust the optimization objective, and make the classification results close to the category of sugarcane, so as to improve the classification accuracy and the classification effect.

2. Dataset processing analysis

2.1 Data situation

This paper by geographical conditions monitoring data as sample data set, data sets and 20 images data, image data for the Beijing no. 2, the multispectral data, image acquisition time for October 2018, the size of each image data set is 5000 pixels by 5000 pixels, geographical conditions monitoring data by artificial markers for images, the sugar cane figure on the spot to drew capture part of the sample are shown in figure 1 below.Figure 1 (a) shows the image sample data, while figure 1 (b) shows the manually labeled sugarcane plot, the green one is the planting range of sugarcane, and the black one is the planting range of non-sugarcane.
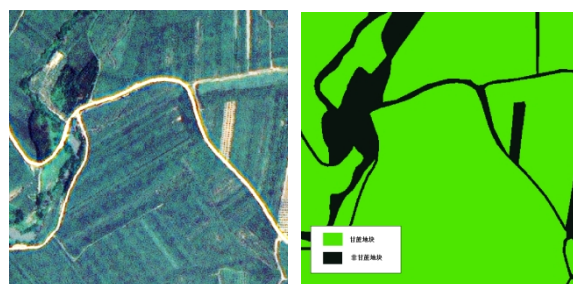
(a)Original image   (b)Cane Ground truth

Figure 1. Remote Sensing Image And GT

2.2 Analysis and processing

Prior to training, data should be pre-processed to solve the following problems: 1. Data imbalance, 2. Data outliers, 3.After preprocessing the remote sensing image data, it can be used as the training data to conduct model training. The following is mainly to conduct statistics on the existing sample data set and make statistics on the number of positive samples.
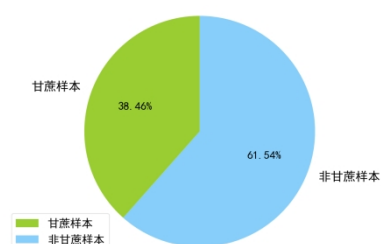

Figure 2. Sugarcane sample statistics

As can be seen from the statistical chart, the statistical number of sugarcane samples accounts for 38.46% and the proportion of non-sugarcane is 61.54%. The number of sugarcane samples is relatively small and the proportion of negative samples is relatively large.2. Keep the number of negative samples unchanged and increase the number of positive samples.In this paper, the second method is adopted to keep the number of negative samples unchanged and increase the number of positive samples through data enhancement so as to adjust the proportion of positive and negative samples.

Due to differences in equipment, weather and terrain conditions, aerial remote sensing images taken have problems such as haze, color deviation, shadow and different colors among multiple images. The quality of these images is a direct factor affecting the overall visual effect and interpretation effect of orthographic correction images(Li, 2013).This is because before making the samples, the remote sensing image samples need to be homogenized. After the homogenization of all the image samples, the quality of the remote sensing image is guaranteed.
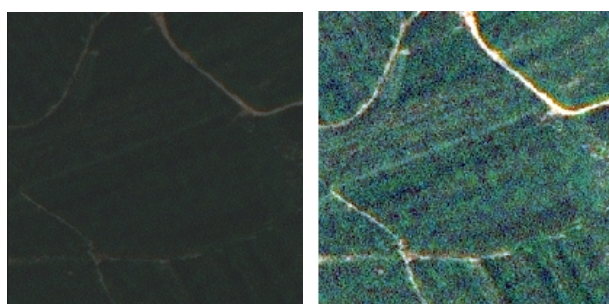
(a)Uniform color   (b)Uneven color

Figure 3. Uniform color

Deep learning all the time since there is a large amount of calculation, easy model fitting problem, in order to solve the problem of over fitting, smooth solution can be trained from the model to solve the skills or the model itself of [], at the same time also can begin through data, data enhanced by data in hand, to enhance the existing data method, so as to adapt to the effect of different characteristics.Data enhanced methods mainly include: random cutting, from top to bottom and turn around, color dithering, adding noise, rotation, translation, scaling, affine transformation, etc., suitable for remote sensing image enhancement method mainly has: rotation, translation, scaling, due to the remote sensing images are classified based on texture feature, so the color dithering is not applicable, use color dithering data to enhance the opposite.

As shown in figure 5 (a) as the original image data, (b) as the data after adding noise, (c) after the flip horizontal data, from the point of data, using data augmented after the texture characteristics of original image doesn't change much, but increased the diversity of samples, to prevent a fitting model training.
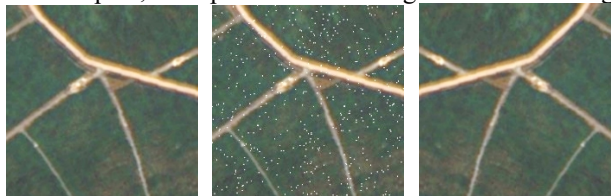
(a)Image   (b)Add Noise   (c)Flip horizontal

Figure 4. Data augmentation comparison

After data enhancement of the original training samples, the diversity of samples was enriched and the balance of samples was better. The pie chart below shows the statistical proportions of sugarcane samples and non-sugarcane samples after data enlargement.
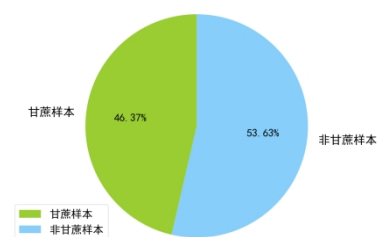

Figure 5. After Data augmentation Sugarcane sample statistics

3 Lovász hinge

Lovasz hinge loss function is a hinge loss function extended based on Lovasz, which is mainly aimed at the situation of dichotomy.

3.1 Jaccard loss

Jaccard index is mainly used to calculate the similarity between samples, which is A statistic used to compare the similarity and diversity of sample sets(Yu J, Blaschko M,2015). For sets A and B, it is defined as the proportional relationship between the intersection size of two sets and the size of the union. The formula is defined as:

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{1}$$

When A and B overlap completely，J(A,B)=1，so J(A,B) for the range of：$0 \leq J(A,B) \leq 1$，The larger the value of J(A,B), the more similar the two sets are, and therefore can be used 1-J(A,B) as a Jaccard loss，The definition for：

$$\Delta_{J_c}(A,B) = 1 - J(A,B) \tag{2}$$

In semantic segmentation, the truth value is labeled as Y^*,the he predicted result is$\tilde{Y}$，so Jaccard index is defined as：

$$J_c(Y^*,\tilde{Y}) = \frac{|\{Y^* = c\} \cap \{\tilde{Y} = c\}|}{|\{Y^* = c\} \cup \{\tilde{Y} = c\}|} \tag{3}$$

Jaccard loss is defined as：

$$\Delta_{J_c}(Y^*,\tilde{Y}) = 1 - J_c(Y^*,\tilde{Y}) \tag{4}$$

The above formula can be rewritten as：

$$M_c(Y^*,\tilde{Y}) = \{Y^* = c, \tilde{Y} \neq c\} \cup \{Y^* \neq c, \tilde{Y} = c\} \tag{5}$$

3.2 Lovász extension

Lovasz extension is a very useful structure to minimize submodular functions(Berman M , Triki A R,2017). The expansion of Lovasz extension is always convex, which can effectively minimize submodular functions，Therefore, the minimum value of Lovasz extension can be used to find the minimum value of submodular functions(Yu J, Blaschko M,2015).

Have set functions $\Delta:\{0,1\}^p \to \mathbb{R}$，with $\Delta(0)=0$，so Lovász extension is defined as(Yu J, Blaschko M,2015)：

$$\overline{\Delta}: m \in \mathbb{R}^p \mapsto \sum_{i=1}^{p} m_i g_i(m) \qquad (6)$$

$$\text{with } g_i(m) = \Delta(\{\pi_1...\pi_i\}) - \Delta(\{\pi_1...\pi_{i-1}\}) \quad (7)$$

3.3 Lovász hinge

In machine learning, hinge loss, as a loss function, is usually used in maximun-margin algorithm(Li, 2012). Hinge loss formula is:

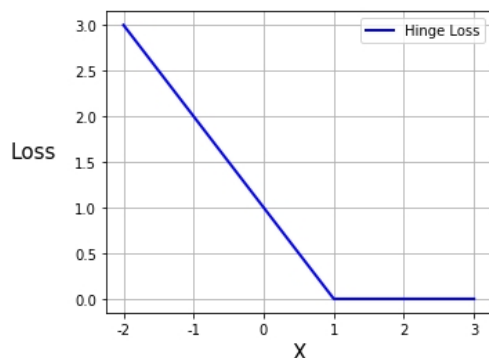$$\ell(y) = \max(0, 1 - y \cdot \hat{y}) \qquad (8)$$

Figure 6. Hinge Loss function curve

Hinge loss function based on Lovasz expansion is Lovasz hinge, which is mainly aimed at dichotomous tasks. After expansion based on Lovasz, the formula is:

$$\hat{l}(s) = \max_{\pi} \sum_{j=i}^{p} s^{\pi_j} \big(l(\{\pi_1,...,\pi_j\}) - l(\{\pi_1,...,\pi_{j-1}\}) \big)$$

4 Experiment

This experiment used multi-spectral data from Beijing no. 2 with a resolution of 0.8 meters. The growth cycle of sugar cane was trained using image data from October 2018.The number of training iterations was 100, the optimizer used SGD stochastic gradient descent, and the accuracy of the final experimental results was 91.67%, which was 2.23% higher than 89.42% without Lovasz hinge.

4.1 Model

Segmentation model based on Net U - is a u-shaped structure network, the model is used in coding and decoding structure with jump connection, adopted convolutional coding structure plus maximum pool to extract image low-level feature value(ZHANG, 2015), on the decoding structure using the sampling image features, in the process of decoding on sampling used jump structure will be connected with the resolution of the encoder output.

After the u-net network structure was proposed in the MICCAI conference in 2015, it achieved good results in various segmentation tasks(Ronneberger O, Fischer P,2015), especially in the beginning, it made a leap breakthrough in medical image segmentation.
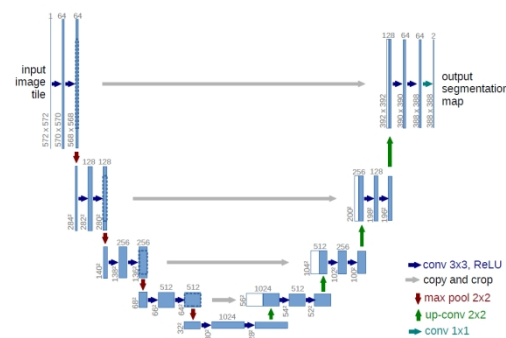

Figure 7. Unet Model(Ronneberger, 2015)

4.2 Loss Function

This experiment adopted the method of comparison. Firstly, cross entropy was used as a loss function for model training. Secondly, Lovasz hinge loss function was added into the comparison experiment to form a combined loss function with cross entropy for model training：

$$L_c = -[y\log \hat{y} + (1-y)\log(1-\hat{y}) \qquad (9)$$

The L_c said cross entropy loss function，y said ground truth value,， $\hat{y}$ said of the predicted value.

The formula of the combined loss function is：

$$L = L_c + 0.5 * l \qquad (10)$$

The L_c said cross entropy loss function，l said Lovász hinge loss function，The loss function of the

final combination is 0.5times Lovász hinge add the cross entropy loss function.

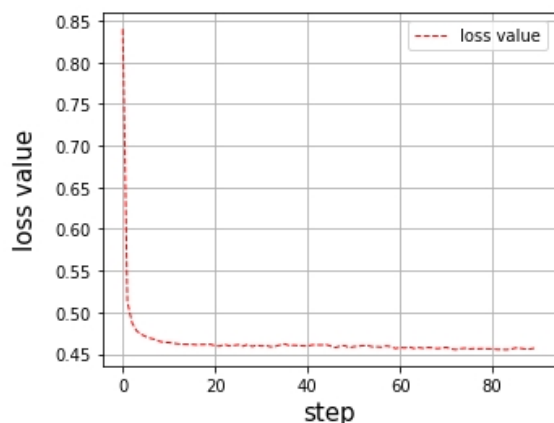Finally, the combined loss curve after training is shown as follows：

Figure 8. Cross entropy curve

5 Results analysis

It can be seen from the experimental comparison that the loss function Lovasz hinge based on Lovasz expansion has a better effect, and the marginal part has a better classification effect after joining Lovasz hinge, as shown in the figure below.
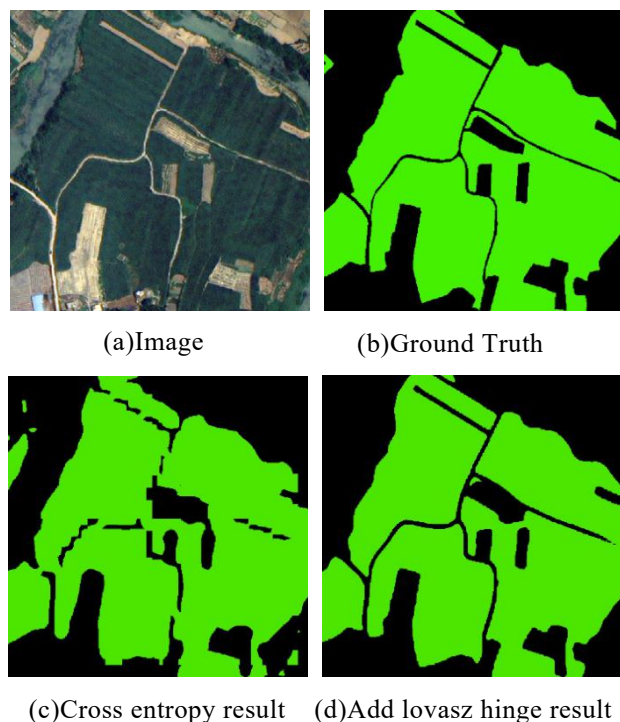

(a)Image   (b)Ground Truth

(c)Cross entropy result   (d)Add lovasz hinge result

Figure 9. Comparison of experimental results

Model Loss Function	Accuracy	IoU	Kappa
Cross Entropy	0.8942	0.7462	0.7780
Lovász hinge	0.9167	0.7696	0.8034

Table 1. Accuracy contrast chart

After joining Lovasz hinge, the training was carried out, and the verification accuracy of the training process was shown in the figure below.
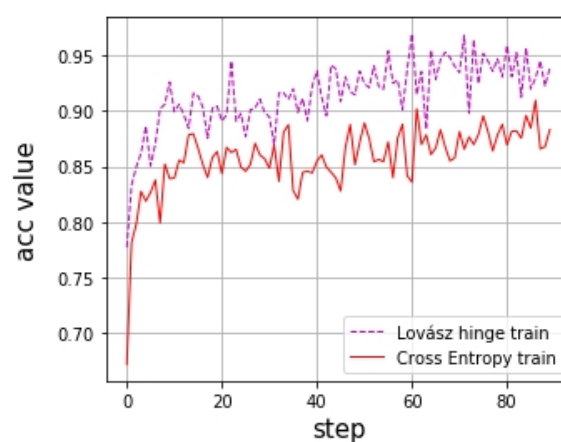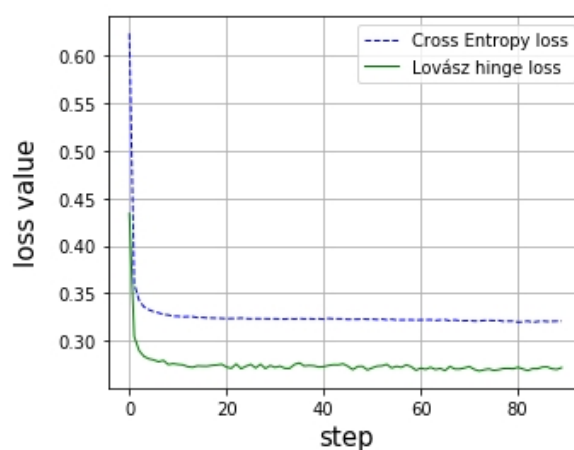

Figure 10. Accuracy curve

Figure 11. Loss train curve

According to the analysis, in the classification task of sugarcane in remote sensing image, the simple Lovasz hinge loss function has a lower value than the cross entropy function. By combining the loss function of Lovasz hinge and cross entropy, the sugarcane classification can be extracted more effectively.The training accuracy is two percentage points higher than that of the cross entropy loss function

6 Conclusion

In this paper, the loss function was mainly used for optimization and adjustment, and Lovasz hinge loss function was added on the basis of the original cross entropy loss function to achieve better optimization effect by combining the loss function. After repeated experiments, 0.5 times of Lovasz hinge loss plus cross entropy was obtained to obtain the optimal experimental effect.

References

ZHANG Liangepei, WU Chen. Advance and Future Development of Chage Detection for Multi-temporal Remote Sensing Imagery[J]. Acta Geodaetica et Cartographica Sinica, 2017(10):249-261.

WU Guangming, CHEN Qi, Shibasaki R , et al. High Precision Building Detection from Aerial Imagery Using a U-Net Like Convolutional Architecture[J]. Acta Geodaetica et Cartographica Sinica, 2018, v.47(06):178-186.

Li li. The process flow Research and Application of unifier ray and color for the aerial remote sensing image[D]. 2013.

Yu J, Blaschko M. The Lovász Hinge: A Convex Surrogate for Submodular Losses[J]. 2015.

Berman M , Triki A R , Blaschko M B . The Lovász-Softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks[J]. 2017.

Li hang. Statistical learning methods [M]. Tsinghua university press, 2012.

ZHANG HuaBo, Research and application of image segmentation by deep learning[D]. School of information and software engineering, 2018.

Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[C] International Conference on Medical Image Computing & Computer-assisted Intervention. 2015.