Application of Random Region Augmentation Algorithm in Deep Learning

Tao Chen1*, Hongying Lu2, Sihe Xiao3

1College of Computer Science and Cyber Security (Oxford Brookes College), Chengdu University of Technology, Chendu Sichuan 610059, China
2College of Computer Science and Cyber Security (Oxford Brookes College), Chengdu University of Technology, Chendu Sichuan 610059, China
3Chengdu University of Technology, Chendu Sichuan 610059, China
*Corresponding author’s e-mail: 417414695@qq.com

Abstract. In the field of computer vision, the collection and sorting of image data is the core driving force. However, the current data collection work cannot perfectly collect the image data of each actual landing scene. The purpose of the data augmentation algorithm is to increase the diversity of the data set and improve the robustness of the model. Traditional data augmentation methods include geometric augmentation and color augmentation, mainly including flipping, rotating, cropping, translation, stretching, zooming, adding noise, blurring, Dropout, Cutout, color jittering. Traditional data augmentation methods have certain limitations, and the effect is not obvious. Based on the idea of Cutout algorithm, this paper proposes the RRA augmentation algorithm, which divides four quadrant regions in the image, and randomly selects the ROI region in each region, and is different from the Cutout algorithm directly discarding the region, but randomizing the region Enhance the color, and finally do geometric augmentation processing on the overall image. Compared with the original single data augmentation operation, the algorithm improves precision by 7%, and recall improves by 7%.

1. Introduction
After entering the 21st century, the rapid development of the Internet has caused a surge in bitstream data in the network, which also announced the advent of the era of big data. Taking this opportunity, the wave of artificial intelligence disciplines has also begun to sweep the world. Recently, machine learning technology has begun to gradually move towards deep learning [1] with the help of large-scale data and high-performance GPU. In computer vision, traditional data augmentation techniques are generally divided into geometric augmentation type and color augmentation type. Geometric augmentation is mainly based on the original image of the image flip [2], rotation [3], cropping [4], translation, stretching, zooming [5-6] and other affine transformations [7], etc. The color augmented types mainly include adding noise [8], blur processing, Dropout [9], Cutout [10] and color jittering. Traditional data augmentation usually processes a single sample. Although, the diversity of sample features can be increased, the improvement space is limited. In response to this bottleneck, subsequent scholars proposed various data augmentation methods for multiple samples. Based on the Cutout method, this paper embeds other traditional geometric augmentation and color augmentation, and proposes a random area augmentation RRA algorithm, which effectively improves the precision and recall of the model.
2. Traditional data augmentation algorithm

2.1 Geometry-augmented

Flip can also be divided into vertical flip and horizontal flip. Vertical flip means that the image will vertically swap the upper and lower pixels along a line that crosses the midpoint of the image height and is parallel to the X axis; horizontal flip means that the image will be located along a line in the image The width center, and the line parallel to the Y axis exchanges the left and right pixels horizontally.

Rotation means that the image takes the center point as the axis of rotation, and the pixels are rotated one or more times in a clockwise or counterclockwise angle, and finally the part that exceeds the original image is discarded. After the movement, the area that is not filled by other pixels is replaced by a constant. The constant is generally (0,0,0) or (0), that is, black pixels.

2.2 Color-augmented

Adding noise refers to the separate processing of the values of random pixels in the original image. Commonly used noises include: Gaussian noise, Laplace noise, Poisson noise, impulse noise, and salt and pepper noise.

Blur processing refers to processing pixel values within a certain range. Commonly used blurs include: Gaussian blur, mean blur, median blur, and two-way blur. The blur processing method takes into account that when the image is collected, the image information in the distance is often less and the image information is very blurred, or some low-end devices have insufficient resolution of their own collection and cause the image to be blurred.

Color jittering refers to the random change of the exposure, saturation and hue of the image to form pictures under different lighting and colors to achieve the purpose of data augmentation, as far as possible to enable the model to use different lighting conditions and improve the generalization ability of the model.

Dropout refers to the loss or noise processing of image information at random pixels of the image.
This method is borrowed from the Dropout processing of neurons in the neural network, and can enhance the model's feature extraction of the overall information of the target image. The CoarseDropout proposed later is to achieve conversion by losing information on a rectangular area with a selectable area and a random location.

Cutout refers to discarding, constant value filling or noise filling in a random rectangular area of the image. This area is generally within the size range of 0.2 to 0.3 of the original image. Generally, 2 to 4 cutout operations are performed that allow overlapping areas.

3 Random regional augmentation algorithm

The traditional Dropout augmentation algorithm is to fill a random single pixel of the image with a value of 0. The subsequent CoarseDropout expands from the original single pixel to a certain area, and can also be discarded at the channel layer. However, the direct discarding method erases the original feature information too much, and it is easy to cause the training to be difficult to converge, or the training to converge slowly.

The random regional augmentation (RRA) algorithm draws on the idea of the Cutout augmentation algorithm, and embeds the traditional geometric augmentation and color augmentation into the ROI area, and no simple deletion processing is performed on the ROI area.

The RRA algorithm step is to divide the original image into four quadrants, then randomly select the ROI area from the four quadrants, and process these areas with 2~3 random color augmentation algorithms, and finally perform 1~2 times on the overall image Random geometric augmentation algorithm processing.

This method combines the special processing of the local features of the image by the Cutout algorithm, and no longer uses the traditional direct discarding method, but uses a softer color augmentation method to enhance the selected area, while retaining the important features of the image, it can also effectively help the model to identify and extract the feature, and improve the convergence speed and detection effect of the model.

4 Experiment and result analysis

Experimental environment: the programming language uses python3.7.8 version, the deep learning framework uses torch1.2.0, the operating system is linux, the CPU model is Intel Core i5-9400, and the GPU model is NVIDIA GeForce GTX 1050 Ti.

The experimental comparison objects are: no data augmentation, geometric data augmentation, color data augmentation, Dropout series data augmentation, and RRA data augmentation. Training uses resnet34[11] as the basic classification network. After 100 epochs are trained, the Accuracy, Precision, Recall and F1-score in the training set and validation set are counted.

| Experiment and result analysis | Accuracy | Precision | Recall | f1-score | ROC | AP |
|--------------------------------|----------|-----------|--------|----------|-----|----|
| Agument free                   | 0.43     | 0.47      | 0.43   | 0.42     | 0.89| 0.40|
| Geometry augmentation          | 0.66     | 0.68      | 0.66   | 0.66     | 0.97| 0.67|
| Color augmentation             | 0.34     | 0.40      | 0.35   | 0.35     | 0.86| 0.31|
| Cutout augmentation            | 0.43     | 0.48      | 0.43   | 0.42     | 0.90| 0.41|
| RRA augmentation               | 0.73     | 0.75      | 0.73   | 0.73     | 0.97| 0.71|

From the table of comparative experimental results in Table 1, it can be seen that geometric augmentation greatly improves the accuracy of the model. It can be seen that the use of appropriate augmentation methods to process data can effectively enhance the diversity of the data set, thereby improving the robustness of the model. After using color data augmentation to enhance the data, the
effect of the model does not increase but decreases. It is speculated that the color augmentation is the overall processing of the image. On a small sample data set, it may cause serious loss of the target feature information, thereby affecting the training of the model effect. After the Cutout data augmentation method is used, the improvement effect of the model is not obvious. After using RRA data augmentation, we can see that the effect is the best of all augmentations. RRA data augmentation continues the advantages of geometric augmentation, and is different from the extreme processing method of color augmentation, but performs a small range of colors in the ROI area.

Figure 5. confusion matrix of RRA data augmentation

It can be seen from the confusion matrix without data augmentation and RRA data augmentation that the RRA data augmentation algorithm can greatly improve the accuracy of each category. The ROC and P-R curve graphs augmented by RRA data also reflect that the data augmentation algorithm has a good performance on the model's Precision and Recall performance indicators.

Figure 6. ROC curve of RRA data augmentation
5 Conclusion
The RRA algorithm refers to the idea of extracting the ROI area in the Cutout algorithm. It discards the Cutout algorithm and randomly selects the ROI processing method for the image as a whole. Instead, it uses the method of dividing the quadrants and selecting the ROI in the quadrants, so that the ROI area can be evenly distributed in the entire image middle. Subsequently, the color augmentation processing is incorporated, which limits the extremeness of the color augmentation algorithm to a large extent, but at the same time retains the characteristics of the color augmentation algorithm. Finally, the overall image is processed by the geometric augmentation algorithm, which inherits the excellent performance of the geometric augmentation algorithm. Experiments have proved that the algorithm has improved precision by 7% and recall by 7% compared to the original single data augmentation operation, which can effectively improve the diversity of the data set and improve the robustness of the model.

References
[1] Hinton G E, Osindero S, Teh Y W. A Fast Learning Algorithm for Deep Belief Nets[J]. Neural Computation, 2014, 18(7):1527-1554.
[2] Xu Pengfei. The realization of several geometric transformations in image processing[J]. Computer Knowledge and Technology, 2017, 13(35):210-212.
[3] Yuan Fenggang, Liu Jiancheng. Study on digital image rotation with different interpolation methods[J]. Software Guide, 2010, 9(04):187-189.
[4] Jiang Hongcheng. Research on automatic cropping, mosaic and color balance of multiple remote sensing images[D]. Graduate University of Chinese Academy of Sciences (Institute of Remote Sensing Applications), 2004.
[5] Wang Sen, Yang Kejian. Research and Implementation of Image Scaling Algorithm Based on Bilinear Interpolation[J]. Automation Technology and Application, 2008(07):44-45+35.
[6] Zhang Yang. Design and implementation of image scaling algorithm based on bilinear interpolation[J]. Electronic Design Engineering, 2016, 24(03):169-170+175.
[7] Zeng Wenfeng, Li Shushan, Wang Jiangan. Translation, rotation and scaling in image registration based on affine transformation model[J]. Infrared and Laser Engineering, 2001(01):18-20+17.
[8] Sun Haiying. Research on denoising algorithm of image Gaussian noise and salt and pepper noise[D]. Fudan University, 2012.
[9] Krizhevsky A, Sutskever I, Hinton G. ImageNet Classification with Deep Convolutional Neural
Networks[J]. Advances in neural information processing systems, 2012, 25(2).

[10] Devries T, Taylor G W. Improved Regularization of Convolutional Neural Networks with Cutout[J]. 2017.

[11] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/ CVPR.2016.90.