Remote Sensing Image Retrieval using Convolutional Neural Network Feature

Dilsha Mol P K[1], Angel P Mathew[2]

[1]*Student, Dept. of Electronics and communication, Ilahia College of Engineering and Technology, Muvattupuzha, Kerala, India*
[2]*Professor, Dept. of Electronics and communication, Ilahia College of Engineering and Technology, Muvattupuzha, Kerala, India*

---***---

Abstract - *Remote sensing image retrieval (RSIR) is a crucial undertaking in remote detecting. Most substance based RSIR approaches accept a basic separation as likeness criteria. A recovery strategy dependent on weighted separation and fundamental highlights of Convolutional neural network (CNN) is proposed in this letter. The strategy contains two phases. In the first place, in disconnected stage, the pre prepared CNN is calibrated by some marked pictures from the objective informational collection, at that point used to remove CNN includes, and named the pictures in the recovery informational collection. Second, in online stage, we utilize the adjusted CNN model to separate the CNN highlight of the inquiry picture and compute the heaviness of each picture class and apply them to figure the separation between the question picture and the recovered pictures. Investigations are led on two RSIR informational indexes. Contrasted and the best in class techniques, the proposed strategy is rearranged yet productive, fundamentally improving recovery execution.*

Key Words: **CNN, RSIR**

1. INTRODUCTION

With the advancement of remote detecting innovation, remote detecting pictures become significant in the field of land examination, urban arranging, catastrophic event observing and appraisal, climate expectation, asset examination, etc. Step by step instructions to consequently and effectively recover the remote detecting pictures that clients need from enormous picture databases gets one of the difficult and rising exploration subjects in the field of remote detecting.

Content based image retrieval (CBIR) is a helpful technique to take care of the difficult issue, which depends on the highlights of picture, for example, force, shape, surface, and structure. In this manner, CBIR has been received by the remote detecting network for improving remote sensing image retrieval (RSIR). Content based RSIR (CBRSIR) is a functioning and testing research point that has pulled in the consideration of scientists around the globe.

Ordinary CBRSIR strategies predominantly utilize low-level highlights to speak to the substance of the remote detecting pictures, which incorporates surface highlights, shading (ghastly) highlights, and shape highlights. The neighborhood, low-level highlights are frequently changed into mid-level portrayals to improve the recovery execution by include encoding strategies, for instance, vector of privately amassed descriptors [1], sack of visual words (BoVW) [2], [3], and pack of-highlights [4].

Motivated by the extraordinary achievement of profound learning in picture arrangement and item location, the remote detecting network has applied profound learning systems to remove significant level semantic highlights for CBRSIR. Ge et al. [5] proposed two collecting Convolutional neural network (CNN) highlights.

One is amassed legitimately by normal pooling with various pooling locale sizes while the other is totaled by BoVW. Two compelling plans, utilizing the pre prepared CNN models and a novel self-planned CNN design, individually, were proposed to remove CNN highlights for high-goals remote sensing image retrieval [6]. Hu et al. [7] presented multi scale link for convolutional highlights and multi fix pooling for completely associated layers to recovery.

The above papers utilized a straightforward separation to ascertain the likeness between the question picture and the recovery picture. This letter utilizes a weighted separation and the straightforward CNN highlights to recover picture. We adjust the pre prepared CNN model to figure the heaviness of each class in the recovered informational collection for question picture. We give more inclination to the recovered pictures in increasingly comparable classes with the inquiry picture.

2. Methodology

In this segment, we first quickly present two CNNs that we use and the calibrating with the two pretrained CNN models, at that point portray the CNN highlights used to recover picture. Next, the weighted separation is introduced. At last, the procedure of the proposed picture recovery strategy is portrayed.

2.1 CNN Models

CNNs have just been viewed as the best profound learning approach in different fields, for example, image classification [8], object acknowledgment [9], [10], and picture scene examination [11]–[13], which comprise of numerous layers, for example, convolution, pooling, and completely associated layers.

Two effective CNN models pretrained on ImageNet are assessed for remote detecting picture recovery in our ongoing work, specifically, the well known gauge model VGG [14] and the ResNet model [15]. VGG model increases great execution without complex calculation, which won the subsequent position in the opposition of ILSVRC in 2014. In 2015, ResNet won the ILSVRC and accomplished cutting edge exhibitions in numerous PC vision errands. ResNet vigorously developed on the convolutional organize by adjusting lingering capacity and character planning.

2.2 Fine-Tuning the CNN Models

It is hard to prepare a novel model from the earliest starting point as a result of the absence of information and hard modification of boundaries. So calibrating is the most ideal approach to show signs of improvement impact with not many emphasess by changing the pretrained boundaries to all the more likely suit the objective informational index.

In this letter, the pretrained CNN models, VGG and ResNet pretrained on ImageNet informational index, are calibrated utilizing the objective recovery informational index. The two pretrained CNNs are actualized also, calibrated utilizing MatConvNet, which is a MATLAB application of CNNs for PC vision applications. MatConvNet gets favorable circumstances in CNN building obstructs under MATLAB which is an inviting advancement condition. MatConvNet makes it conceivable to make and make CNN design by utilizing significant level programming language as opposed to low-level dialects and keep productivity on CPU or GPU calculation.

Concerning calibrating process, the softmax layer standardized by the last completely associated or convolutional layer is utilized for arrangement; in this manner, the quantity of neural units in this layer for the most part equivalents to picture class number of the informational index. Hence, we modify the last completely associated or convolutional layer class number of yields to accommodate our objective informational index class number. The loads of the last layer are haphazardly introduced utilizing a Gaussian appropriation with mean 0 and fluctuation 0.01. The loads are refreshed by versatile second estimation (Adam) advancement calculation with a learning rate of 0.001, a force of 0.9, and a weight rot of 0.0005. We utilize the default information expansion path in Matconvnet to increase information.

2.3 Feature Extraction

The yields from the more significant level layers of CNN (e.g. Fc7 and Fc8 from VGG Model) have been demonstrated to be effective nonexclusive highlights which accomplish cutting edge execution in picture recovery [5]. Here, we consider the yields of significant level layer (Fc7 and Fc8) from the adjusted VGG16 model and the last convolutional layer from the adjusted ResNet50 model as the highlights for remote detecting picture recovery.

2.4 Labeling Image Retrieval Data Set

Each picture in the recovery informational collection should be named which class it has a place with. We feed each picture into the finetuned display and apply the softmax capacity to the yield of the calibrated model for changing over the yield to the likelihood for each class.

2.5 Weighted Distance

CNN is exceptionally successful in recognizing picture class, so we utilize the capacity of CNN to improve remote detecting picture recovery. The inspiration of weighted separation is to give more inclination to the recovered pictures in increasingly comparative classes with the question picture.

The more the class likelihood of the question picture, the less is the separation between the question picture and the recovered pictures of that class. At the point when a question picture q is taken care of into the adjusted model, we get a likelihood pq for each class and use it to compute the heaviness of a recovered picture in the recovery informational index concurring

$w = 1 - p^k_q$

where k is the class of the picture r . The weighted separation between the question picture q and the recovered picture r is reformulated as

$dw(q, r) = w \times d(q, r)$

where d(q, r) is determined utilizing the ordinary separation picked reasonably. The Euclidean separation is utilized to assess the exhibition of the proposed technique.

2.6 Process of Image Retrieval

It comprises of two sections: the disconnected part and the online part. In the disconnected part, we adjust the pretrained CNN model with a few marked pictures. At that point, we name and concentrate the CNN highlights for the pictures in the recovered informational collection with the tweaked CNN model. At long last, we assemble an informational index with include vectors furthermore, class marks.

In the online piece of our technique, a question picture is taken care of into tweaked CNN model to figure the CNN highlights and the class likelihood. At that point, the weighted separations between the question picture and the recovered pictures are processed. From that point onward, we sort the recovered pictures in climbing request as per the weighted separations. At long last, we get the recovery results.

3. EXPERIMENTS AND ANALYSIS

Here, we assess the presentation of the proposed strategy for CBRSIR on two remote detecting informational indexes which are freely accessible. The first is the University of California, Merced (UCMD) informational index [3], which comprises of 21 land-use classes trimmed from the United States Geological Survey National Map. Each class has 100 pictures with 256×256 pixels. The other one is PatternNet [16], which is a huge scope high-goals openly accessible informational index. The information set contains 38 classes, every one of which has 800 pictures with a size of 256×256 pixels gathered from Google Earth or U.S. urban areas scenes.

To approve the adequacy of the proposed strategy, we contrast our strategy and the best in class techniques. Since most related works depend on UCMD, the execution examination on UCMD with the cutting edge techniques is appeared in fig 1. The aftereffects of our strategies are better than those of different strategies when the quantity of preparing pictures abundances 10. The most minimal normal standardized adjusted recovery rank worth is only 0.04 with the proposed techniques. In quickly, our technique is improved and can get amazing execution with some preparation pictures.

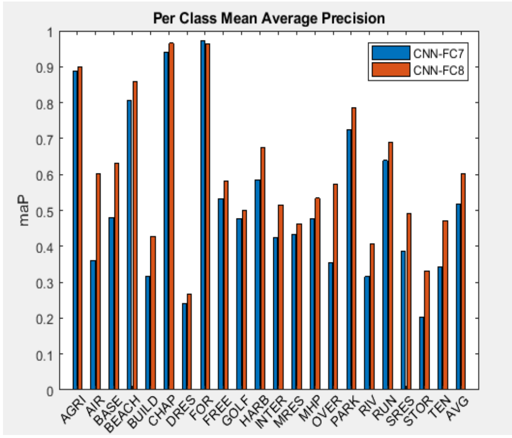

Chart – 1: Comparison with previous method

Fig - 1 : Input query image

Fig - 2 : Output retrieval image

4. CONCLUSIONS

In this letter, we introduced a remote detecting picture recovery strategy dependent on weighted separation and CNN. In the disconnected stage, we utilized the adjusted CNN models to extricate picture highlights and name the class of picture in the recovery informational index.

In the online stage, we determined the heaviness of each class as indicated by the class likelihood of the inquiry picture and utilized it to alter the separation between the inquiry picture furthermore, the recovered pictures. The exploratory outcomes on the UCMD and PatternNet informational collections show that the proposed technique accomplishes better execution contrasted and that of the best in class strategies.

REFERENCES

[1] S. Özkan, T. Ates, E. Tola, M. Soysal, and E. Esen, "Performance analysis of state-of-the-art representation methods for geographical image retrieval and categorization," IEEE Geosci. Remote Sens. Lett., vol. 11, no. 11, pp. 1996–2000, Nov. 2014.

[2] Y. Yang and S. Newsam, "Geographic image retrieval using local invariant features," IEEE Trans. Geosci. Remote Sens., vol. 51, no. 2, pp. 818–832, Feb. 2013.

[3] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in Proc. ACM Int. Conf. Adv. Geogr. Inf. Syst., 2010, pp. 270–279.

[4] N. Passalis and A. Tefas, "Learning neural bag-of-features for largescale image retrieval," IEEE Trans. Syst., Man, Cybern. Syst., vol. 47, no. 10, pp. 2641–2652, Oct. 2017.

[5] Y. Ge, S. Jiang, F. Ye, C. Jiang, Y. Chen, and Y. Tang, "Aggregating CNN features for remote sensing image retrieval," Remote Sens. Land Resour., to be published.

[6] W. Zhou, S. Newsam, C. Li, and Z. Shao, "Learning low dimensional convolutional neural networks for high-resolution remote sensing image retrieval," Remote Sens., vol. l9, no. 5, p. 489, May 2017.

[7] F. Hu, X. Tong, G.-S. Xia, and L. Zhang, "Delving into deep representation for remote sensing image retrieval," in Proc. Int. Conf. Image Process., Nov. 2017, pp. 198–203.

[8] Y. Yu and F. Liu, "Aerial scene classification via multilevel fusion based on deep convolutional neural networks," IEEE Geosci. Remote Sens. Lett., vol. 15, no. 2, pp. 287–291, Feb. 2018.

[9] W. Min, M. Fan, X. Guo, and Q. Han, "A new approach to track multiple vehicles with the combination of robust detection and two classifiers,"IEEE Trans. Intell. Transp. Syst., vol. 19, no. 1, ppl 174–186, Jan. 2018.

[10] Q. Wang, J. Gao, and Y. Yuan, "Embedding structured contour and location prior in siamesed fully convolutional networks for road detection," IEEE Trans. Intell. Transp. Syst., vol. 19, no. 1, pp. 230–241, Jan. 2018.

[11] W. Min, H. Cui, H. Rao, Z. Li, and L. Yao, "Detection of human falls on furniture using scene analysis based on deep learning and activity characteristics," IEEE Access, vol. 6, pp. 9324–9335, Jan. 2018.

[12] Q. Wang, J. Gao, and Y. Yuan, "A joint convolutional neural networks and context transfer for street scenes labeling," IEEE Trans. Intell. Transp. Syst., vol. 19, no. 5, pp. 1457–1470, May 2018.

[13] Q. Wang, F. Zhang, and X. Li, "Optimal clustering framework for hyperspectral band selection," IEEE Trans. Geosci. Remote Sens., to be published, doi: 10.1109/TGRS.2018.2828161.

[14] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in Proc. Int. Conf. Learn. Represent., Sep. 2015, pp. 1–14.

[15] Z. Wu, C. Shen, and A. van den Hengel. (2016). "Wider or deeper:Revisiting the ResNet model for visual recognition."[Online].Available:https://arxiv.org/abs/1611.10080

[16] W. Zhou, S. Newsam, C. Li, and Z. Shao, "PatternNet: A benchmark dataset for performance evaluation of remote sensing image retrieval,"ISPRS J. Photogramm. Remote Sens., to be published, doi: 10.1016/j.isprsjprs.2018.01.004.

[17] E. Aptoula, "Remote sensing image retrieval with global morphologicaltexture descriptors," IEEE Trans. Geosci. Remote Sens., vol. 52, no. 5,pp. 3023–3034, May 2014.

[18] [18] Z. Du, X. Li, and X. Lu, "Local structure learning in high resolution remote sensing image retrieval," Neurocomputing, vol. 207, pp. 813–822, Sep. 2016.