Journal of Information and Computing Science
J. Info. Comput. Sci., Vol. 19, No. 1, pp. 65-80

DOI: 10.4208/JICS-2024-005
June 2024

Remote Sensing Image Scene Classification Based on Deep Learning Feature Fusion

Liqi Wang[1], Cheng Zhang[1], Yuchao Hou[2], Xiuhui Tan[1], Rong Cheng[1], Xiang Gao[1] and Yanping Bai[1,2,*]

[1] *School of Mathematics, North University of China, Taiyuan 030051, China*

[2] *School of Information and Communication Engineering, North University of China, Taiyuan 030051, China*

Abstract. In view that traditional manual feature extraction method cannot effectively extract the overall deep image information, a new method of scene classification based on deep learning feature fusion is proposed for remote sensing images. First, the Grey Level Co-occurrence Matrix (GLCM) and Local Binary Patterns (LBP) are used to extract the shallow information of texture features with relevant spatial characteristics and local texture features as well; second, the deep information of images is extracted by the AlexNet migration learning network, and a 256-dimensional fully connected layer is added as feature output while the last fully connected layer is removed; and the two features are adaptively integrated, then the remote sensing images are classified and identified by the Grid Search optimized Support Vector Machine (GS-SVM). The experimental results on 21 types of target data of the public dataset UC Merced and 7 types of target data of RSSCN7 produced average accuracy rates of 94.77% and 93.79%, respectively, showing that the proposed method can effectively improve the classification accuracy of remote sensing image scenes.

AMS subject classifications: 68U10, 68T05

Key words: Image classification, Convolutional Neural Network (CNN), Grey Level Co-occurrence Matrix (GLCM), Local Binary Patterns (LBP), migration learning, Support Vector Machine (SVM)

---

1 Introduction

Translated from *Journal of Nanjing University of Information Science & Technology*, 2023, 15(3): 346-356.
*Corresponding author. *Email addresses:* **1023252901@qq.com** (L. Wang), **baiyp666@163.com** (Y. Bai).

With the continuous development of remote sensing technology [1], remote sensing image classification has been widely applied in many fields such as land management, urban planning and traffic supervision [2]. However, at present, remote sensing scene images contain rich and complex information and structures, and there are still many challenges in how to make reasonable use of rich information in remote sensing images to obtain accurate and effective features [3].

Traditional manual feature extraction is commonly used in remote sensing image scene classification, including color histogram, texture feature, Global Image Descriptor (GIST), Scale Invariant Feature Transform (SIFT), etc. Li Fuyu et al. [4] pointed out that the remote sensing image registration technology based on SIFT has advantages in scale rotation invariance. Xu Junyi et al. [5] took Grey Level Co-occurrence Matrix (GLCM) as the first principal component of Principal Component Analysis (PCA), making full use of the robustness of GLCM in acquiring texture features. Although traditional manual features have good stability and the ability to express the overall shallow information, and are feasible to be directly applied to the scene classification task of low-resolution remote sensing images, traditional manual features are too dependent on manual design and cannot effectively extract the feature information of high-resolution remote sensing images, which makes them not widely applied in classification tasks.

In order to effectively solve the above problems and the lack of generalization ability and poor classification performance caused by a single feature, scholars have proposed a variety of feature fusion classification methods. Chen Xu et al. [6] proposed a texture classification algorithm based on the fusion features of GLCM and Tamura, which enhanced the robustness and classification performance of the algorithm by improving the rotation invariance of GLCM and reducing a large amount of redundant information. Zhang Qingchun et al. [7] used a multi-feature fusion algorithm to extract local entropy, texture features and other features to improve image classification performance. Wang Yu et al. [8] adopted a new spatial feature-the fusion of second-order moment feature and spectral feature-to achieve road refinement. Kang Jian et al. [9] used the RFB module to obtain high-level water body semantic information and multi-scale, integrated initial multi-scale features with original features in a deep level, completed the extraction of multi-scale features, and enhanced high-level water body semantic information features. These methods not only consider the global feature information, but also retain the shallow local information. Through the fusion of shallow information and global feature information, the generalization ability and classification performance of the algorithm are improved to a certain extent. However, this kind of feature fusion will increase the amount of computation, which leads to the increase of model complexity and the occurrence of overfitting. Therefore, PCA module is introduced in this algorithm, and appropriate principal component contribution rate is selected to remove redundant information and improve the classification performance of the model.

In recent years, due to the superiority of deep learning algorithm in image recognition, many scholars have introduced Convolutional Neural Network (CNN) into remote sensing image scene classification. Although CNN has achieved good results in