Performance Analysis of Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models for Land cover Image Classification: A Comparative Study

Cham Kotage

10 May 2023

1 Introduction

Accurately categorizing different types of land cover from image data presents a significant challenge due to the complex and diverse nature of land surfaces. This study addresses the problem of land cover classification using machine learning algorithms, specifically focusing on the performance of two approaches: Support Vector Machine (SVM) and Convolutional Neural Network (CNN). A comparative analysis of these models was conducted for an image classification task involving cloud images. The CNN model achieved a test accuracy of 0.0286, whereas the SVM model attained a higher test accuracy of 0.0571. Based on these results, it can be concluded that the SVM model outperformed the CNN model on this particular dataset.

2 Method

2.1 Data preparation

The classification tasks for each model involved several common stages, including data preprocessing and the generation of training, validation, and testing sets. Once the images and their corresponding labels were loaded from the UC Merced Land Use Dataset and checked for label-image consistency, the dataset was shuffled to mitigate potential biases associated with data ordering.The images were subsequently normalized to enhance the model's learning process. This was achieved by dividing the pixel values of the images by 255, resulting in a scaling of the pixel values within the range of 0-1. The Land Use dataset consisted of 21 distinct classes, each containing 100 images. After normalization, the dataset was divided into training, validation, and testing subsets, ensuring an ample amount of data for model training, separate data for hyperparameter tuning (validation), and an independent set for the final evaluation of model performance (testing). The generated sets were then used as input for the SVM and CNN models to perform the classification task.

2.2 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are widely employed in image classification research due to their deep learning capabilities[SJM18]. The CNN model operates by accepting input in the form of images, subjecting them to multiple convolutional and pooling layers to extract distinctive features, and subsequently utilizing dense layers to make accurate predictions pertaining to each image class.

For configuring the CNN model, the input shape was determined based on the training set. A convolutional layer with 32 filters of size 3x3 and ReLU activation was established to capture local image features such as edges and shapes. A max pooling layer downsamples the feature maps to prevent overfitting. A flatten layer is then added to prepare the data for the fully connected layer. The fully connected layer integrates low-level features from the convolutional layers and converts them into abstract and task-specific representations, enhancing the precision of land cover classification.

The CNN model was then compiled before training. The model's compilation employed the adam optimizer and sparse categorical cross-entropy loss. The adam optimizer is a favorable selection for its effectiveness and minimal memory demands[MPV19]. A kernel size of (3,3) exhibits satisfactory performance while avoiding an escalation in computational complexity for the model. The 'fit' function was employed to train the model, incorporating the training, testing, and validation sets specified in section 2.1 as input parameters. During each epoch of the training process, the model acquires the ability to identify patterns and features within the images corresponding to various land cover categories. Regular updates on the model's training progress are provided at each epoch to facilitate the monitoring of its performance.

After model training, performance was assessed by evaluating accuracy across the training, validation, and testing datasets. Predictions were generated using the 'predict' method to determine class names for each dataset. Accuracy was calculated using the 'accuracy score' function from scikit-learn. Additionally, a confusion matrix compared predicted and original classes, providing insights into correct and incorrect predictions for each class. A visualization of the confusion matrix was created using the Matplotlib library.

2.3 Support Vector Machine

Support vector machines are commonly used machine learning algorithms used in classification and regression tasks. These models are well known in literature for their capability to handle non-linearly separable data [WFW21]. In this particular task, SVMs are utilized to perform the classification by finding an optimal hyperplane that effectively separates different classes within the feature

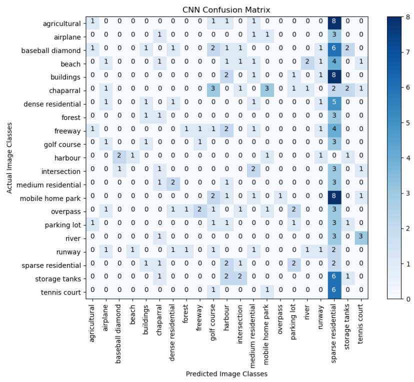

Figure 1: CNN Confusion Matrix

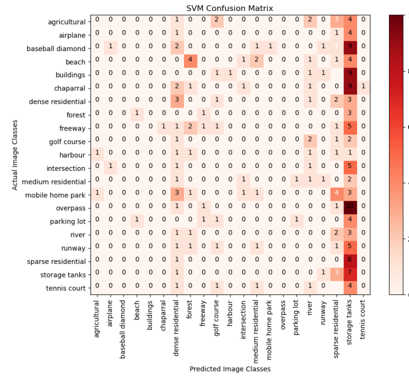

Figure 2: SVM Confusion Matrix

space. The SVM classifier was instantiated using the 'svm' library from scikit-learn to initiate the model production. Feature extraction was performed using the Histogram of Gradients (HOG) feature descriptor, which transformed the image data into a representation capable of capturing relevant visual patterns for land cover classification.

The defined SVM model was trained on the training set, which consisted of the extracted features and their corresponding labels, using the 'fit' method. During the training process, the SVM learns the ideal decision boundary that separates the different land cover categories. By maximizing the margin between different classes, the decision boundary allows for improved generalization and robustness. Once trained, the SVM model is employed to predict labels for the validation set features using the 'predict' method. The predicted labels are then compared to the actual labels to evaluate the model's performance on unseen data. The calculated test accuracy provides a final assessment of the model's ability to accurately classify new land cover images.Similar to the steps described in Section 2.2, another confusion matrix was generated to gain insights into the model's performance on the test set. This matrix provides an overview of the predicted and actual classes for each image, facilitating an assessment of both correct and incorrect predictions made by the model.

3 Results

The confusion matrix serves as a valuable visual tool to assessing the performance of a given model across various classes, aiding in the identification of classes that are particularly challenging for classification. Within the matrix, the diagonal cells correspond to correct predictions, where the predicted class aligns with the actual class. Conversely, the off-diagonal cells represent incorrect predictions. In Figure 1, the confusion matrix generated for the CNN model is presented, while Figure 2 exhibits the confusion matrix for the SVM model.

After examining the confusion matrix of the CNN (Figure 1), it becomes evident that the vertical row corresponding to the predicted label 'spare residential' exhibits a notable concentration of dark-colored cells. This observation suggests that the CNN model frequently assigns this particular image class to various actual classes. Hence, it is evident that the model possesses a bias towards predicting the 'spare residential' class. Consequently, numerous misclassifications across different classes lead to a noticeably low precision for the 'spare residential' class.

In line with Figure 1, the confusion matrix produced by the SVM model (Figure 2) demonstrates a similar pattern. Notably, the row associated with the predicted land cover category 'storage tanks' displays a substantial concentration of cells shaded in dark color, indicating a significant occurrence of misclassifications. This finding implies a consistent misidentification tendency of the SVM model, where diverse land cover classes are frequently mistaken as 'storage tanks'. The discernible misclassification pattern suggests that the features utilized by the classifier exhibit resemblances between the 'storage tanks' class and other classes, ultimately resulting in inaccurate predictions.

4 Conclusion

By comparing the two models, we can identify the pros and cons of implementing them in land cover classification. The CNN model had a tendency to misclassify instances in the 'sparse residential' class, while the SVM model showed similar misclassification patterns for instances labeled as 'storage tanks'. These models have distinct biases, prioritizing different aspects of land cover images during classification. The apparent issue observed in both of the proposed image classification models is the presence of a consistent misclassification pattern, wherein numerous images belonging to different classes are erroneously categorized as either 'storage tanks' or 'sparse residential' classes.

Further analysis is required to address the observed preference for a specific class and enhance the classification performance of the proposed models. Specifically, focusing on misclassified instances categorized as 'storage tanks' or 'spare residential' could reveal common patterns contributing to the confusion. The convolutional layers of CNNs excel in capturing spatial hierarchies in image data, while support vector machines (SVMs) aim to identify optimal hyperplanes for class separation. Analyzing these discrepancies may shed light on their preferences for individual classes.

It is important to acknowledge that the classification of the 'sparse residential' and 'storage tanks' categories may pose challenges for all classifiers, not just the two that we explored in this paper. To enhance the feature extraction process and refine the classification approach, it is advisable to involve domain experts who can provide insights into the distinctive characteristics of these misclassified land cover classes. Furthermore, it is recommended to explore additional data preprocessing and augmentation techniques to enhance the performance of the classifiers in these specific cases

5 References

sources

References

[SJM18] Neha Sharma, Vibhor Jain, and Anju Mishra. "An Analysis Of Convolutional Neural Networks For Image Classification". In: *Procedia Computer Science* 132 (2018). International Conference on Computational Intelligence and Data Science, pp. 377–384. ISSN: 1877-0509. DOI: `https://doi.org/10.1016/j.procs.2018.05.198`. URL: `https://www.sciencedirect.com/science/article/pii/S1877050918309335`.

[MPV19] Smit Mehta, Chirag Paunwala, and Bhaumik Vaidya. "CNN based Traffic Sign Classification using Adam Optimizer". In: *2019 International Conference on Intelligent Computing and Control Systems (ICCS)*. 2019, pp. 1293–1298. DOI: `10.1109/ICCS45141.2019.9065537`.

[WFW21] Pin Wang, En Fan, and Peng Wang. "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning". In: *Pattern Recognition Letters* 141 (2021), pp. 61–67. ISSN: 0167-8655. DOI: `https://doi.org/10.1016/j.patrec.2020.07.042`. URL: `https://www.sciencedirect.com/science/article/pii/S0167865520302981`.

6 Appendix