Applied GIS

ISSN: 1832-5505

Vol-13 Issue-02 April 2025

Autonomous Landing Scene Recognition Based on Transfer Learning for Drones

[1]M. Sravanthi, [2]Ungur Sravani, [3]Kataray Sri Neha, [4]Polu Susmitha

[1] Assistant Professor, Department of Information Technology, Bhoj Reddy Engineering College for Women, Vinay Nagar, Hyderabad

[2,3,4] Student, Department of Information Technology, Bhoj Reddy Engineering College for Women, Vinay Nagar, Hyderabad

Abstract

In this work, we investigate drone autonomous landing scene detection via knowledge transfer. The challenges associated with aerial remote sensing—namely, the fact that various images have distinct representations at different altitudes or that some pictures are very similar—led us to use a deep convolutional neural network (CNN) that is based on knowledge transfer and fine-tuning to address the issue. Next, the seven classes comprise the LandingScenes-7 dataset is created. Furthermore, we use thresholding in the prediction step to take care of the classifier's ongoing novelty detection issue by excluding additional landing scenes. The adaptive momentum (ADAM) optimization technique is used in conjunction with the ResNeXt-50 backbone to facilitate our transfer learning approach. We also compare momentum stochastic gradient descent (SGD) optimizer with ResNet-50 backbone. ResNeXt-50, which uses the ADAM optimization method, performs better, according to the experiment findings. It is possible for drones to autonomously learn landing scenes using this pre-trained model and fine-tuning, as it achieves 97.8450% top-1 accuracy on the LandingScenes-7 dataset.

Keywords: Drone autonomous landing, scene detection, aerial remote sensing, knowledge transfer

Introduction

Our study delves into the autonomous landing scene identification of drones using knowledge transfer. We used a deep convolutional neural network (CNN) that is based on knowledge transfer and fine-tuning to address the issues related to aerial remote sensing, specifically the fact that different images have distinct representations at different altitudes or that some pictures are very similar. The LandingScenes-7 dataset is then constructed, consisting of seven classes. Moreover, we address the classifier's persistent novelty detection problem by removing extra landing scenes via thresholding in the prediction stage. To support our transfer learning strategy, we combine the ResNeXt-50 backbone with the adaptive momentum (ADAM) optimization technique. Additionally, we contrast the ResNet-50 backbone with the momentum stochastic gradient descent (SGD) optimizer. The experiment results show that ResNeXt-50 works better and employs the ADAM optimization approach. Given that our pre-trained model achieves 97.8450% top-1 accuracy on the LandingScenes-7 dataset, drones may be able to independently learn landing scenes with some fine-tuning.

Related Work

Target categorization in remote sensing photos using an efficient distributed convolutional neural network architecture and pre-training

It is becoming more difficult to identify objects with similar looks using remote sensing images (RSIs) in an effective and efficient manner. Convolutional neural networks (CNNs) are now the dominant method for classifying targets because of their superior performance and strong feature representation capabilities. However, CNN relies mostly on a single computer for testing and training. Because processing RSIs requires a lot of time and limited hardware resources, a single system naturally has limitations and becomes a bottleneck. Furthermore, because of the imbalance between the model structure and the RSI data, the CNN model faces the problem of overfitting. Overfitting happens and results in poor prediction performance when a model is complicated or the training data is small. In order to tackle these issues, a distributed CNN architecture is suggested for RSIs target categorization, which significantly boosts the system's scalability and CNN's training performance. It enhances RSIs' processing effectiveness and storage capacity. Additionally, the CNN model is made more flexible and resilient by using the Bayesian regularization strategy to initialize the CNN extractor's weights. It assists in avoiding local optima brought on by inadequate RSI training pictures or an improper CNN structure, as well as overfitting. Furthermore, taking into account the effectiveness of the Naïve Bayes classifier, a distributed Naïve Bayes classifier is engineered to minimize the training expenses. The suggested system and approach work the best and improve recognition accuracy when compared to other algorithms. The results demonstrate that the suggested algorithms and distributed system architecture are appropriate for target categorization tasks in RSIs.

Challenges, Approaches, Benchmarks, and Opportunities in Remote Sensing Image Scene Classification Combined with Deep Learning

With a wide variety of applications, remote sensing image scene classification seeks to assign a set of semantic categories to remote sensing pictures based on their contents. Deep neural networks' potent feature learning capabilities have propelled the field of remote sensing picture scene categorization, which has garnered notable interest and yielded noteworthy advancements. Nonetheless, as far as we are aware, there hasn't been a thorough examination of current developments in deep learning for remote sensing picture scene categorization. This article offers a comprehensive overview of deep learning techniques for remote sensing picture scene categorization, including over 160 publications, in light of the field's rapid advancement. Specifically, we go over the three main approaches to remote sensing image scene classification and survey challenges: autoencoder-based, convolutional neural network-based, and generative adversarial network-based. Furthermore, we provide an overview of the benchmarks used in remote sensing picture scene categorization and provide a performance summary of over two dozen sample methods on three widely-used benchmark datasets. We conclude by talking about the exciting prospects for further study.

Using the ™-channel to improve object recognition

We suggest adding a new channel, known as the ™-channel, to traditional RGB photos and using it for a variety of classification and identification tasks. With a cheap frosted glass put in front of one of the binocular cameras, a binocular camera is used to concurrently record the same scene in the color and frosted light channels, as seen in the new RGB-TM picture. The frosted glass's ability to scatter light results in an imperfect frosted light path. In this work, we provide a unique optimization to optimize the ℓ-channel to retain edges owing to scene radiance, directed by the RGB channel. Our RGB-TM pictures are effective, as shown by extensive testing findings that show significant gains in a range of scene categorization and object identification tasks.

Locations: An Image Database of 10 Millions for Scene Identification

Data-hungry machine learning algorithms can already do almost human-level semantic categorization on tasks like visual object and scene identification thanks to the growth of multi-million-item dataset efforts. We provide here the Places Database, a collection of 10 million scene photos annotated with scene semantic categories that provide a comprehensive and varied inventory of the many kinds of surroundings found around the globe. We present scene classification CNNs (Places-CNNs) as baselines using the most advanced Convolutional Neural Networks (CNNs), which perform noticeably better than the earlier methods. Upon seeing the CNNs trained on Places, it becomes evident that object detectors function as a middle-tier representation for scene categorization. Together with Places-CNNs, the Places Database provides a fresh resource to steer future work on scene identification difficulties because to its great variety and coverage of exemplars.

SUN Database: Examining an Extensive Range of Scene Categories

Understanding the complex and varied visual settings that comprise our everyday lives is necessary for progress in scene comprehension. In pursuit of this goal, we provide the Scene Understanding database, an almost complete set of scenes classified with the same degree of detail as spoken language. There are 131,072 photos in the database, organized into 908 different scene types. We do a thorough study of co-occurrence statistics and the contextual link using this data, which has both scene and item labels accessible. In two human tests, we assess the accuracy of human scene identification and analyze the typicality of each picture within its assigned scene category to get a deeper understanding of this extensive taxonomy of scene categories. We next carry out three computational experiments: "scene detection," where we loosen the assumption that a single picture represents a single scene category, indoor vs outdoor scene categorization, and scene identification using global image attributes. The link between human and machine recognition mistakes as well as the relationship between picture "typicality" and machine recognition accuracy are finally explored, and we compare the results of human studies to the performance of machines.

Adam: An Approach to Probabilistic Optimization

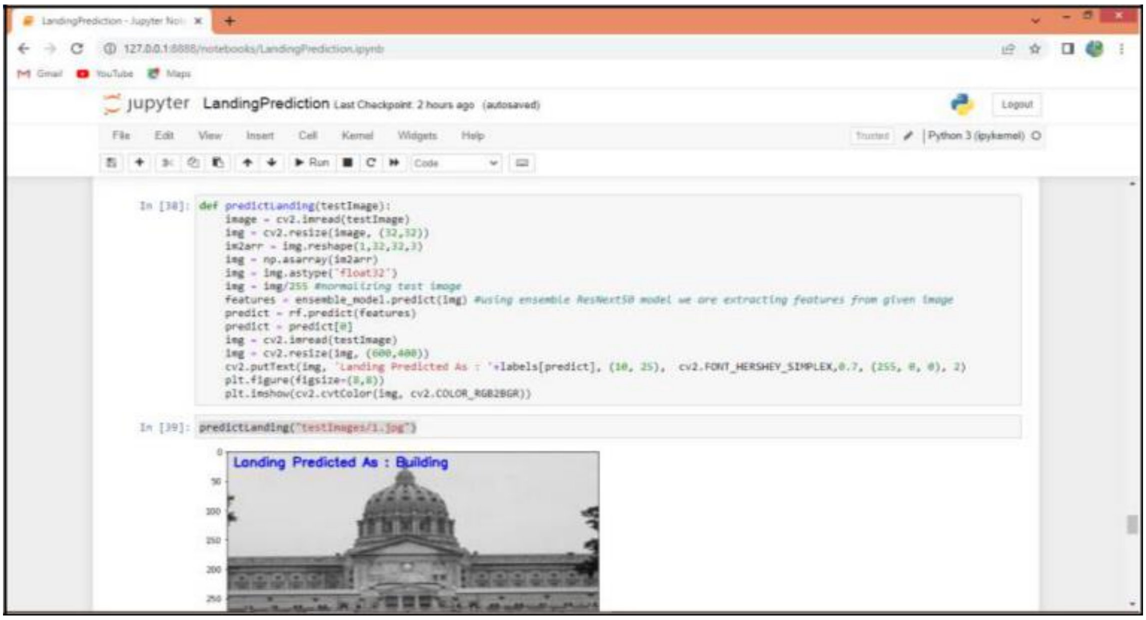
We present Adam, an adaptive lower-order moment estimator-based technique for first-order gradient-based optimization of stochastic objective functions. Large problems with lots of data and/or parameters are a good fit for this approach since it is easy to construct, computationally efficient, requires minimal memory, and is invariant to diagonal rescaling of the gradients. Non-stationary targets and issues with very noisy or sparse gradients may also benefit from this approach. The hyper-parameters are usually easily interpreted and don't need to be adjusted too much. There is a discussion of some links to similar algorithms that Adam was influenced by. Additionally, we examine the algorithm's theoretical convergence features and provide a regret constraint on the convergence rate that is on par with the most well-known outcomes under the online convex optimization framework. Empirical findings show that Adam performs well in real-world scenarios and holds its own against other stochastic optimization techniques. In conclusion, we address Ada Max, an Adam variation grounded in the infinite norm.

Methodology

1. **Get label:** Using this module we will get the label
2. **calculate metrics:** Using this module, metrics can be calculated
3. **Predict landing:** Using this module we will predict the landing position

Results and Discussion

The training graph for the ResNext50 is shown above. The x-axis shows the training epoch, and the y-axis shows the accuracy and loss values. The accuracy line is represented



by a green line, while the loss line is represented by a blue line. As the epoch progressed, the accuracy increased and approached 1, while the loss decreased.

The predict function is defined in the above graph. It takes an input picture path and, using an extended ensemble object, classifies the provided image scenes; in the example scene above, it is classified as a building.

Conclusion

The authors of this research investigate drone knowledge transfer for autonomous landing scene detection. We use a deep convolution neural network (CNN) based on knowledge transfer and fine-tuning to address the issue, taking into account the challenges in aerial remote sensing, particularly the fact that some pictures are very similar or the same scene has distinct representations at various altitudes. Next, a dataset called LandingScenes-7 is created and classified into seven classifications. Additionally, the classifier still struggles with novelty detection, which we solve by eliminating additional landing scenes using the thresholding method in the prediction step. Using the adaptive momentum (ADAM) optimization technique, we use the transfer learning approach based on the ResNeXt-50 backbone. The momentum stochastic gradient descent (SGD) optimizer and the ResNet-50 backbone are also contrasted.

References

1. Li BQ, Hu XH. Effective distributed convolutional neural network architecture for remote sensing images target classification with a pre-training approach. J Syst Eng Electron. 2019;30(2):238-244. doi:10.21629/JSEE.2019.02.12.
2. Xia GS, Bai X, Ding J, *et al*. DOTA: a large-scale dataset for object detection in aerial images. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition doi:10.1109/CVPR.2018.00417.
3. Cheng G, Xie XX, Han JW, *et al*. Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities. IEEE J Sel Top Appl Earth Obs Remote Sens. 2020;13:3735-3756. doi:10.1109/JSTARS.2020.3002216.
4. Anand AK, Barman S, Prakash NS, *et al*. Vision based automatic landing of unmanned aerial vehicle. In: Proceedings of the International Conference on Information Technology and Applied Mathematics. 2019:102-113. doi:10.1007/978-3-030-48392-2_12.
5. Tian C, Huang CB. An algorithm for unmanned aerial vehicle landing scene recognition based on deep learning and computational verbs. In: Proceedings of the IEEE International Conference on Civil Aviation Safety and Information Technology. 2019:180-184. doi:10.1109/ICCAVSIT.2019.8945648.
6. Lu CW, Tsougenis E, Tang C. Improving object recognition with the l-channel. Pattern Recognit. 2016;49:187-197. doi:10.1016/j.patcog.2015.09.029.
7. Zhou BL, Lapedriza A, Khosla A, *et al*. Places: a 10 million image database for scene recognition. IEEE Trans Pattern Anal Mach Intell. 2018;40(6):1452-1464. doi:10.1109/TPAMI.2017.2723009.
8. Xiao JX, Hays J, Ehinger KA, *et al*. SUN database: large-scale scene recognition from abbey to zoo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2010:3485-3492. doi:10.1109/CVPR.2010.5539970.
9. Xiao JX, Ehinger KA, Hays J, *et al*. SUN database: exploring a large collection of scene categories. Int J Comput Vis. 2016;119(1):3-22. doi:10.1007/s11263-015-0815-8.
10. Patterson G, Xu C, Su H, *et al*. The SUN attribute database: beyond categories for deeper scene understanding. Int J Comput Vis. 2014;108(1-2):59-81. doi:10.1007/s11263-013-0695-z.

ISSN: 1832-5505
Vol-13 Issue-02 April 2025

11. Yang T, Li PQ, Zhang HM, *et al*. Monocular vision SLAM-based UAV autonomous landing in emergencies and unknown environments. Electronics. 2018;7(5):73. doi:10.3390/electronics7050073.
12. He KM, Zhang XY, Ren SQ, *et al*. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-778. doi:10.1109/CVPR.2016.90.
13. Xie SN, Girshick R, Dollar P, *et al*. Aggregated residual transformations for deep neural networks. In: Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2017:5987-5995. doi:10.1109/CVPR.2017.634.
14. Xie L, Lee FF, Liu L, *et al*. Scene recognition: a comprehensive survey. Pattern Recognit. 2020;102:107205. doi:10.1016/j.patcog.2020.107205.