The International Journal of Interdisciplinary Organizational Studies
ISSN: 2324-7649 (Print), ISSN: 2324-7657 (Online)
Volume 19, Issue 1, January-June, 2024

AUTONOMOUS LANDING SCENE RECOGNITION BASED ON HYBRID ENSEMBLE RANDOM FOREST MODEL

M. Raghavendra Reddy[1], Dr. C. Anbu Ananth[2]
[1]Assistant Professor, Department of CSE, G Pulla Reddy Engineering College (Autonomous), Kurnool, AP.
[2]Associate Professor, Department of CSE, Chidambaram, Annamalai University, Annamalai Nagar, TN.
1raghureddy224@gmail.com, 2anbu_ananth2006@yahoo.com

Abstract
This study focuses on drones' recognition capabilities of their landing scenes during data transfer. To tackle the problems of aerial remote sensing, including highly similar images or scenes with different representations at different heights, we employ a deep convolutional neural network (CNN) that is based on knowledge transfer and fine-tuning. This results in the creation of the Landing Scenes-7 dataset, which is divided into seven categories. Moreover, we use the thresholding technique to eliminate more landing scenes during the prediction stage, which resolves the classifier's ongoing novelty detection problem. We apply a transfer learning technique using the Hybrid Ensemble Random Forest model and the ResNeXt-50 backbone. We also assess the momentum stochastic gradient descent (SGD) optimizer and the ResNet-50 backbone in conjunction with it. The experimental results show that ResNeXt-50 using the ADAM optimisation approach performs better. With a pre-trained model and some fine-tuning, drones will soon be able to learn landing scenes on their own; this model obtained 97.8450% top-1 accuracy on the LandingScenes-7 dataset.

INDRODUCTION

Based on the various platforms used for operation, remote sensing can be classified as ground, aerial, or aerospace. Sensors installed on drones, helicopters, and balloons are used for aerial remote sensing. The advancement and utilisation of convolutional neural networks (CNNs) and graphics processing units (GPUs) has made it feasible for drones to identify landing scenes. Due to the recent emergence of CNN and public datasets, significant advancements have been made in object recognition in aerial photos and scene classification in remote sensing images. These include geographical object detection, scene classification of satellite images, and aerial scene classification. Deep learning and the label "H" were used by Anand et al. to accomplish autonomous drone landing. Nevertheless, the landing mark might not be set up beforehand in emergency landing situations. Tian et al. used the Inception V3 model to identify landing scenes on the ImageNet dataset. To increase the accuracy of the recognition, they suggested a learning rate decay technique using computational verb theory. An approach was presented by Lu et al. to enhance the l-channel of conventional RGB photographs. The outcomes of the experiments demonstrate the usefulness of the l-channel, as evidenced by the notable enhancement in performance in tasks involving object recognition and scene classification. The use of datasets like Places365, Sconce Understanding (SUN), and SUN attribute datasets is essential for CNN-based scene recognition. In an emergency situation or unfamiliar location, Yang et al. presented a simultaneous localization and mapping (SLAM) system based on a monocular vision to enable autonomous landing for drones. He and his team won the ImageNet large scale visual recognition challenge (ILSVRC)-2015 and COCO-2015 competitions with their clever backbone, Reset. They also won the tasks of common object in context (COCO) detection and segmentation, ImageNet detection, and localization. Xie et al. demonstrated a ResNeXt-50 with group = 32 and bottleneck width d = 4 using group convolution that Reset inherited. Nevertheless, the network outperformed its Reset counterparts only in the cases where a building block was "bottleneck." It performs better, specifically, on ResNeXt-50, ResNeXt-101, or ResNeXt-152. Numerous well-liked and successful methods for scene recognition were examined by Xie et al. While a large body of literature has been written about scene recognition in the past, not much research has been done on drone landing scenario classifications. We must confront the distinctive representation of the landing scene in order to obtain recognition. An region may be designated as a "water area" rather than a "wilderness," for instance, if there are several lotus leaves floating on the lake.

An area of "water area" or "road" might be used to describe a scene in which a bridge crosses the Yangtze River. Stated differently, landing scene recognition is more concerned with the background of the image, whereas object detection concentrates on the image's foreground. Recognition of landing scene types is one of the difficult difficulties. Realising perfect autonomy and intelligence for drones has always been the aim of robotics research. Drones that follow the intended path and avoid obstacles on their own are currently capable of autonomous flight. On non-emergency circumstances, nevertheless, these autonomous flights operate. A drone must determine whether the present scene below it is appropriate for an emergency landing in the event of a rapid reduction in battery power or malfunction. Consequently, this paper's study aim is to investigate landing scene identification appropriate for drones. The drone's flying safety can be enhanced in this instance through landing scene recognition.

This work makes the following significant contributions: The network is first initialised using pre-trained weights, then we retrain the model using fine-tuning, all based on a knowledge transfer technique using the Resent. The second dataset developed is called LandingScenes-7, and it contains photographs from Places365, the Internet, and drone aerial photography. We partition the dataset into seven categories: water area, wilderness, heatfield, road, lawn, and vehicle-intensive site. Using the dataset LandingScenes-7, we first train the CNN backbone-based model via transfer learning. Through the novelty detection module, we secondly determine whether the test stage result is accurate. If the likelihood of a certain category is high, it is output directly; if not, thresholding is used to make additional decisions.

Related work

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

It is becoming increasingly difficult to identify targets 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 machine 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, the imbalance between the model structure and the RSI data makes overfitting a problem for the CNN model. Overfitting happens and results in poor predictive performance when a model is complex 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 speed. It enhances RSIs' processing effectiveness and storage capacity. Additionally, the CNN model is made more flexible and robust by using the Bayesian regularisation approach to initialise the CNN extractor's weights. By avoiding local optima brought on by insufficient RSI training pictures or an improper CNN structure, it aids in preventing overfitting. Furthermore, taking into account the effectiveness of the Naïve Bayes classifier, a distributed Naïve Bayes classifier is engineered to minimise the training expenses. The suggested system and method work the best and improve recognition accuracy when compared to other algorithms. The outcomes 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 range of applications, remote sensing image scene classification seeks to assign a set of semantic categories to remote sensing images based on their contents. Deep neural networks' potent feature learning capabilities have propelled the field of remote sensing image scene classification, which has garnered notable interest and yielded noteworthy advancements. Nonetheless, as far as we are aware, there hasn't been a thorough examination of recent developments in deep learning for remote sensing image scene classification. This article offers a comprehensive overview of deep learning techniques for remote sensing picture scene classification, encompassing over 160 publications, in light of the field's swift advancement. With regard to remote sensing image scene classification and survey, we specifically address the following main challenges: auto encoder-based methods; convolutional neural network-based methods; and generative adversarial network-based methods. Furthermore, we provide an overview of the benchmarks utilised in remote sensing image 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 more 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 recognition tasks. With a cheap frosted glass placed in front of one of the binocular cameras, a binocular camera is used to concurrently record the same scene in the colour and frosted light channels, as shown in the new RGB-TM image. The frosted glass's ability to scatter light results in an imperfect frosted light path. In this work, we present a novel optimisation to optimise the ℓ-channel to retain edges due to scene radiance, directed by the RGB channel. Our RGB-TM images are effective, as evidenced by extensive testing results that show notable gains in a range of scene classification and object recognition tasks.

Locations: A 10-million-picture database for scenario identification

Multi-million-item dataset projects have made it possible for data-hungry machine learning algorithms to achieve performance close to that of humans in semantic categorization for tasks like visual object and scene recognition. In this article, we present the Places Database, a collection of 10 million scene photos that have been annotated with scene semantic categories. The database includes a wide range of environmental kinds that are often seen throughout the world. We present scene classification CNNs (Places-CNNs) using the most advanced Convolutional Neural Networks (CNNs) as baselines, which considerably surpass the earlier methods. Object detectors emerge as an intermediary representation of scene classification, as demonstrated by the visualisation of the CNNs trained on Places. The Places Database, in conjunction with the Places-CNNs, provides a fresh resource to steer future work on scene recognition difficulties due to its large coverage and high diversity 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 present 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, organised into 908 different scene groups. We do a thorough analysis of co-occurrence statistics and the contextual link using this data, which has both scene and item labels available. In two human tests, we assess the accuracy of human scene recognition and measure the typicality of each image within its assigned scene category to gain 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 image represents a single scene category, indoor versus outdoor scene categorization, and scene recognition using global image attributes. The relationship between human and machine recognition errors as well as the relationship between image "typicality" and machine recognition accuracy are finally explored, and we compare the results of human studies to the performance of machines.

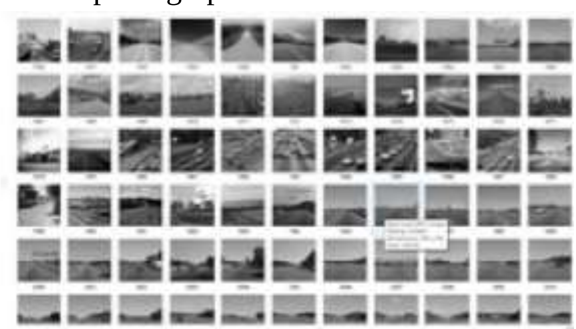
Adam: An Approach to Probabilistic Optimisation

We present Adam, an adaptive lower-order moment estimator-based technique for first-order gradient-based optimisation of stochastic objective functions. Large problems with lots of data and/or parameters are a good fit for this method 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 extremely noisy or sparse gradients can 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 related algorithms that Adam was influenced by. Additionally, we examine the algorithm's theoretical convergence features and offer a regret constraint on the convergence rate that is on par with the most well-known outcomes under the online convex optimisation framework. Empirical findings show that Adam performs well in real-world scenarios and holds its own against alternative stochastic optimisation techniques. We conclude by talking about Ada Max, an Adam variant based on the infinity norm.

Concerning the Dataset

The self-constructed Landing Scenes dataset, which includes eight distinct landing scenarios, was used to train algorithms. The dataset's photos are shown below.

Eight distinct folders are visible in the dataset screen above; simply select any folder to view dataset photographs.



The photographs from the Road class are shown above. We trained models using the location dataset from the above situations, and the drone was then informed of the name of the landing area.

The x-axis in the following graph indicates the names of the landing scenes, while the y-axis shows the number of scene photos that are included in the dataset.

Concerning Algorithms

An Adaptive Group To enhance prediction performance, the Random Forest model combines the ideas of the Random Forest algorithm and ensemble learning. Let's dissect each section:

1. Ensemble Learning: This machine learning paradigm involves training several models, or learners, to answer a single issue. The final prediction is then created by combining the predictions from various models in some fashion, which frequently produces better results than any one model used alone. Techniques like bagging, boosting, and stacking are frequently used in groups.

2. Random Forest: During training, several decision trees are constructed using Random Forest, a well-liked ensemble learning technique that yields the mean prediction (regression) or the mode of the classes (classification). Two main ways that it incorporates randomness are by using random subsets of the data (also known as "bagging") for each tree's training and by only taking into account a random subset of characteristics at each tree split.

A more reliable and accurate predictive model is produced by combining Random Forest with other ensemble approaches or individual models in a Hybrid Ensemble Random Forest model. This is how it could operate:

• **Combining Different Algorithms:** Support vector machines (SVM), gradient boosting machines (GBM), and neural networks are a few examples of base algorithms that can be added to Random Forest in place of just decision trees in the ensemble.

• **Stacking:** This technique entails training several different models and fusing the predictions of those models with a meta-model. One of the foundation models in the stack of a Hybrid Ensemble Random Forest might be Random Forest.

• **Boosting:** Random Forest could be used in conjunction with boosting algorithms like AdaBoost or Gradient Boosting, as an alternative to bagging.

• **Feature Engineering and Selection:** The ensemble's overall performance can be improved by increasing the diversity of the models by incorporating feature engineering or selection strategies.

A hybrid ensemble random forest model's precise implementation can change depending on the particular issue at hand, the properties of the data, and the performance measures that are wanted. The objective is to build a more potent prediction model by utilising the advantages of both Random Forest and ensemble learning.

METHODOLOGY

The modules below are what we designed in order to carry out this project.

1. Importing Classes and Packages: We are importing all necessary Python packages and classes by using this module.

2. Upload Dataset: We will upload the dataset to the application, loop through it, and load all of the photographs from the dataset folder using this module.

3. Dataset pre-processing: We will pre-process the dataset using this module, including normalisation, shuffle, and dataset splitting into train and test subsets.

4. Train Existing ResNext50 Algorithm: We will train the ResNext50 algorithm as transfer learning using this module's CNN layer and ADAM optimizer.

5. Train Proposed ResNext50 Algorithm: We will train the ResNext50 algorithm as transfer learning using this module's CNN layer and ADAM optimizer.

6. Train Hybrid Ensemble Random Forest: Utilising this module will improve our accuracy compared to other methods by training the Extension Hybrid Ensemble Random Forest.

7. Comparison Graph: We will create a comparison graph for each algorithm using this module.

8. Predict Landing: We will anticipate using this module, which will accept an input picture path. It will then classify the given image scenes using an extended ensemble object, and in the example image scene, it will classify the landing type.

RESULTS AND DISCISSION

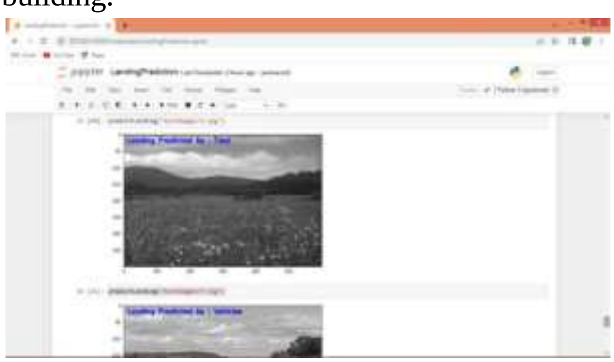
Hybrid Ensemble Random Forest outperformed other algorithms with 100% accuracy in the result shown above.

The x-axis in the comparison graph above denotes the names of the algorithms, the y-axis their accuracy and other metrics in various colour bars, and all of the algorithms' extensions achieved excellent accuracy.

In above screen displaying all algorithm performance in tabular format

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

The graph above defines the predict function. Using an extended ensemble object, it classifies the given image scenes given an input image path; in the example scene above, it is categorised as a building.



Conclusions:

Within this article, we examine autonomous landing scene identification for drones via knowledge transfer. The challenges associated with aerial remote sensing—namely, the fact that different images have distinct representations at different altitudes or that some pictures are remarkably similar—led us to use a deep convolution neural network (CNN) that is based on knowledge transfer and fine-tuning to address the issue. Next, the dataset for landing scenes is created and split up into seven classes. Furthermore, we use thresholding in the prediction stage to take care of the classifier's ongoing novelty detection issue by excluding additional landing scenes. The adaptive momentum (ADAM) optimisation technique is utilised in conjunction with the ResNeXt-50 backbone to facilitate our transfer learning approach. The Hybrid Ensemble Random Forest algorithm was trained by comparing the momentum stochastic gradient descent (SGD) optimizer with the ResNet-50 backbone. The results indicate that this approach yields the best results when compared to other algorithms for autonomous landing scene prediction.

References:

[1] LI B Q, HU X H Effective distributed convolutional neural network architecture for remote sensing images target classification with a pre-training approach
Journal of Systems Engineering and Electronics, 2019, 30 (2): 238- 244.

[2] XIA G S, BAI X, DING J, et al DOTA: a large-scale dataset for object detection in aerial imagesProc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, 3974- 3983.

[3] CHENG G, XIE X X, HAN J W, et al Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities
IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13, 3735- 3756.

[4] ANAND A K, BARMAN S, PRAKASH N S, et al Vision based automatic landing of unmanned aerial vehicleProc. of the International Conference on Information Technology and Applied Mathematics, 2019, 102- 113.

[5] TIAN C, HUANG C B An algorithm for unmanned aerial vehicle landing scene recognition based on deep learning and computational verbs Proc. of the IEEE International Conference on Civil Aviation Safety and Information Technology, 2019, 180- 184.

[6] LU C W, TSOUGENIS E, TANG C Kim proving object recognition with the l-channel
Pattern Recognition, 2016, 49, 187- 197.

[7] ZHOU B L, LAPEDRIZA A, KHOSLA A, et alPlaces: a 10 million image database for scene recognitionIEEE Trans. on Pattern Analysis and Machine Intelligence, 2018, 40 (6): 1452- 1464.

[8] IAO J X, HAYS J, EHINGER K A, et alSUN database: large-scale scene recognition from abbey to zooProc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2010, 3485- 3492.

[9] IAO J X, EHINGER K A, HAYS J, et alSUN database: exploring a large collection of scene categoriesInternational Journal of Computer Vision, 2016, 119 (1): 3- 22.

[10] PATTERSON G, XU C, SU H, et alThe SUN attribute database: beyond categories for deeper scene understandingInternational Journal of Computer Vision, 2014, 108 (1/2): 59- 81.

[11] YANG T, LI P Q, ZHANG H M, et al. Monocular vision SLAM-based UAV autonomous landing in emergencies and unknown environments. Electronics, 2018, 7(5): 73.

[12] HE K M, ZHANG X Y, REN S Q, et alDeep residual learning for image recognitionProc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 770- 778.

[13] IE S N, GIRSHICK R, DOLLAR P, et alAggregated residual transformations for deep neural networksProc. of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2017, 5987- 5995.

[14] IE L, LEE F F, LIU L, et alScene recognition: a comprehensive surveyPattern Recognition, 2020, 102, 107205.