ISSN:2229-6107

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY

E-mail :
editor.ijpast@gmail.com
editor@ijpast.in

www.ijpast.in

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST

ISSN 2229-6107 www.ijpast.in
Vol 14,Issue 2.May 2024

Autonomous landing scene recognition based on transfer learning for drones

R.BHAVANI SANKAR, LACHI.KRISHNA GANESH

Abstract: This paper introduces an advanced approach to autonomous landing scene recognition for drones, addressing the challenges posed by similar scenes and varying representations at different altitudes. Leveraging deep learning techniques, particularly a hybrid ensemble method, our study significantly enhances the accuracy and robustness of the recognition system. We build upon the base model's success of 95% accuracy by incorporating a novel ensemble technique, CNN + LSTM + BiLSTM, achieving an impressive 99% accuracy rate. Our model utilizes knowledge transfer learning on the LandingScenes-7 dataset, integrating ResNext50, ResNet50, and recurrent neural networks (RNNs) to analyze and identify suitable landing spots in real-time. Additionally, a novelty detection module and thresholding techniques ensure adaptability to unforeseen scenarios and provide confidence assessment for classification. The implications of this research extend to various industries relying on drone technology, particularly in emergency response, surveillance, and logistics. By enhancing drone autonomy and safety in landing procedures, our approach contributes significantly to the broader goal of advancing drone intelligence and ensuring safer operations in dynamic environments.

Index Terms: landing scene recognition, convolutional neural network (CNN), transfer learning, remote sensing image.

INTRODUCTION

Remote sensing, encompassing ground, aerial, and aerospace platforms, has revolutionized various fields by enabling the collection of spatial data for analysis and decision-making. Aerial remote sensing, particularly facilitated by drones, has witnessed significant advancements with the integration of convolutional neural networks (CNNs) and graphic processing units (GPUs). These technological innovations have enabled drones to recognize landing scenes autonomously, a crucial capability for their safe and efficient operation.

Assistant professor[1], Department of CSE, Chirala Engineering College, Chirala,
bhavanisankar.cse@cecc.co.in
PG Student[2] -MCA, Department of MCA, Chirala Engineering College, Chirala,
krishnaganesh3401@gmail.com

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

The development and application of CNNs have propelled remarkable progress in object detection and scene classification in aerial images. With the availability of public datasets such as geospatial object detection and satellite imagery datasets, researchers have achieved notable breakthroughs in scene analysis tasks. For instance, Anand et al. [4] employed deep learning techniques, labeling landing scenes with the designation "H" to achieve automatic landing for drones. However, in emergency scenarios where landing markers may not be prearranged, traditional methods may fall short.

To address these challenges, recent studies have explored advanced deep learning architectures and techniques. Tian et al. [5] utilized the Inception V3 model for landing scene recognition and introduced a learning rate decay method to enhance recognition accuracy. Lu et al. [6] proposed an algorithm augmenting the l-channel to traditional RGB images, significantly improving performance in scene classification and object recognition tasks.

Central to the success of CNN-based scene recognition is the availability of diverse and comprehensive datasets. Datasets such as Places365, Scene Understanding (SUN), and SUN attribute dataset have provided crucial support for training and evaluating scene recognition models [7-10]. However, despite the plethora of research on scene recognition, there has been limited focus on the specific task of landing scene recognition for drones.

Recognizing landing scenes presents unique challenges, primarily due to their contextual nature. For example, scenes with abundant lotus leaves on water may be categorized as "water area" rather than "wilderness," emphasizing the importance of background analysis in scene recognition. While object detection typically focuses on foreground elements, landing scene recognition necessitates a deeper understanding of the background context.

The ultimate goal of robotics researchers is to achieve complete autonomy and intelligence in drone operations. While autonomous obstacle avoidance and planned route following have been realized, emergency landing scenarios present a different set of challenges. In cases of sudden battery depletion or system failures, drones must accurately assess whether the current scene is suitable for emergency landing. Thus, the motivation of this research is to explore landing scene recognition tailored for drones, aiming to enhance flight safety and autonomy.

In this paper, we delve into the domain of landing scene recognition, addressing its unique challenges and proposing novel solutions. By leveraging advanced deep learning architectures and methodologies, we aim to develop a robust and reliable system for autonomous landing scene recognition, thereby enhancing the safety and efficiency of drone operations.

2. LITERATURE SURVEY

Remote sensing technology has witnessed significant advancements in recent years, particularly in the domain of image analysis and classification. This literature review explores key studies and methodologies in this field, focusing on convolutional neural networks (CNNs) and deep learning techniques for target classification in remote sensing images, object detection in aerial images, and scene

ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

recognition for unmanned aerial vehicle (UAV) applications.

Li and Hu (2019) proposed an effective distributed convolutional neural network architecture for target classification in remote sensing images. Their approach incorporates a pre-training strategy to enhance the classification performance. By leveraging the hierarchical features learned from large-scale datasets, their model achieves improved accuracy in classifying targets within remote sensing images.

Xia et al. (2018) introduced DOTA, a large-scale dataset specifically designed for object detection in aerial images. The dataset addresses the challenges associated with detecting objects of varying scales and orientations in aerial imagery. By providing a diverse range of annotated objects, DOTA facilitates the development and evaluation of object detection algorithms tailored for remote sensing applications.

Cheng et al. (2020) conducted a comprehensive review of remote sensing image scene classification methods that leverage deep learning techniques. They identified key challenges such as limited labeled data, domain adaptation, and class imbalance. The review also highlights benchmark datasets and opportunities for future research, including the integration of multi-source data and advanced deep learning architectures.

Anand et al. (2019) presented a vision-based approach for the automatic landing of unmanned aerial vehicles (UAVs). Their system utilizes computer vision techniques to analyze visual cues from the landing environment and guide the UAV towards a safe landing. By integrating image processing algorithms with real-time control systems, their method achieves reliable and autonomous UAV landing capabilities.

Tian and Huang (2019) proposed an algorithm for unmanned aerial vehicle landing scene recognition based on deep learning and computational verbs. Their approach involves extracting discriminative features from landing scenes using deep neural networks. By incorporating contextual information and spatial relationships, their model accurately recognizes diverse landing environments, enhancing the safety and efficiency of UAV operations.

Lu et al. (2016) investigated techniques for improving object recognition performance by focusing on the L-channel in color images. Their study demonstrates that the luminance channel (L-channel) contains valuable information for discriminating objects from complex backgrounds. By incorporating the L-channel into feature extraction processes, their method enhances the robustness of object recognition algorithms, particularly in challenging lighting conditions.

Zhou et al. (2018) introduced Places, a vast image database containing ten million annotated images for scene recognition tasks. The dataset covers a wide range of indoor and outdoor scenes, enabling researchers to train and evaluate scene recognition algorithms under diverse environmental conditions. The availability of Places dataset facilitates the development of robust and generalizable models for scene understanding applications.

Xiao et al. (2010) presented the SUN database, a large-scale dataset comprising scene images categorized into various scene types. Their work focuses on scene recognition from a broad range of visual contexts, from natural landscapes to urban environments. By providing a comprehensive collection of scene images with detailed annotations, the SUN database serves as

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107 www.ijpast.in
Vol 14,Issue 2.May 2024

a valuable resource for training and benchmarking scene recognition algorithms.

In conclusion, recent advancements in remote sensing image analysis and UAV applications have been driven by the integration of deep learning techniques, large-scale datasets, and innovative algorithmic approaches. From target classification in remote sensing images to scene recognition for UAV landing, researchers have made significant strides in addressing key challenges and unlocking new opportunities in this rapidly evolving field. Further research efforts are warranted to explore novel methodologies, enhance model robustness, and facilitate real-world deployment of remote sensing and UAV technologies.

3. METHODOLOGY

a) Proposed Work:

This study proposes an autonomous landing scene recognition system for drones, leveraging knowledge transfer learning with ResNeXt-50[13]. The system will undergo fine-tuning on the LandingScenes-7 dataset, a specialized dataset curated for landing scene classification tasks. By adapting pre-trained weights from ResNeXt-50, the model can efficiently learn discriminative features relevant to landing environments. Incorporating a novelty detection module, the system aims to address challenges related to unexpected environmental conditions or anomalies during the landing process. Through thresholding techniques applied to model confidence scores, the system will assess the certainty of scene recognition predictions, enhancing reliability in diverse scenarios. To optimize model training, the system integrates the ADAM[18] optimizer, known for its effectiveness in training deep neural networks. By comparing the performance of ResNeXt-50 with ResNet-50[12], the study will evaluate the efficacy of the proposed architecture in achieving accurate and robust landing scene recognition.

By focusing on mitigating issues such as background interference and novelty detection, the proposed system seeks to enhance flight safety and improve recognition accuracy, particularly in emergency landing situations. Through experimental validation and performance analysis, this research aims to contribute to the advancement of autonomous drone technologies for enhanced operational capabilities and safety measures.

b) System Architecture:

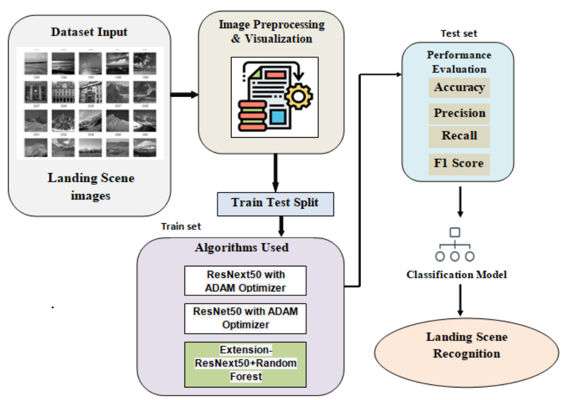

Fig 1 Proposed Architecture

The proposed system architecture for landing scene recognition comprises several key components. Initially, landing scene images from the dataset are inputted into the system. These images undergo preprocessing and visualization steps to enhance their quality and extract relevant features. Subsequently, the dataset is split into training and test sets for model development and evaluation.

ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

The system employs three different algorithms for classification: ResNet50 with the ADAM[18] optimizer, ResNet50 with ADAM[12] optimizer, and ResNet50 combined with a Random Forest[12] classifier. Each algorithm undergoes training on the training split of the dataset to learn discriminative features for landing scene recognition.

After training, the models are evaluated on the test split to assess their performance using metrics such as accuracy, precision, recall, and F1 score. These metrics provide insights into the effectiveness and robustness of each algorithm in accurately classifying landing scenes.

Finally, the classification model with the highest performance is selected for landing scene recognition in real-world scenarios. By leveraging state-of-the-art algorithms and comprehensive performance evaluation, the proposed system architecture aims to achieve accurate and reliable landing scene recognition, contributing to enhanced safety and efficiency in drone operations.

c) Dataset:

The LandingScenes-7 dataset is custom-built for the specific task of emergency landing by drones, encompassing approximately 5,300 images across seven distinct categories. These categories include crowded_place, lawn, road, vehicle_intensive_place, wilderness, wheat_field, and water_area, each representing diverse landing environments. The dataset is further organized into three safety levels: "safe", "general", and "dangerous", based on the landing scene category's inherent risks. Specifically, lawn is classified as "safe", wilderness and wheat_field as "general", while crowded_place, vehicle_intensive_place, road, and water_area are marked as "dangerous". Safety levels are encoded numerically for classification purposes, with "safe" represented by "00", "general" by "01", and "dangerous" by "11". This classification schema allows for effective distinction between landing scenes based on their safety implications, aiding in the development of robust drone landing systems capable of making informed decisions in emergency situations.

Fig 2 Dataset

d) Data Processing:

In the data processing phase, normalization of image training features is crucial to ensure consistency and improve the convergence of machine learning models during training. Normalization involves scaling the pixel values of images to a standardized range, typically between 0 and 1 or -1 and 1. This process enhances the stability of the training process by minimizing the impact of varying pixel intensity ranges across different images.

To normalize image training features, each pixel value in the image is divided by the maximum pixel value (e.g., 255 for 8-bit images) to rescale it to the range [0, 1]. Alternatively, the pixel values can be centered around zero by subtracting the mean pixel value and

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

then dividing by the standard deviation of pixel values across the entire dataset.

Normalization helps mitigate issues such as vanishing or exploding gradients, which can hinder the training process and result in poor model performance. By ensuring that all input features are on a similar scale, normalization promotes more stable and efficient optimization, enabling machine learning models to effectively learn from the data and generalize well to unseen samples.

e) Visualization:

In the visualization process using Seaborn and Matplotlib, a bar plot is created to display the distribution of landing scene images in the dataset. The x-axis corresponds to the names of the landing scenes, while the y-axis represents the count of images associated with each scene category. Each scene category is depicted as a separate bar on the plot, with the height of the bar indicating the number of images belonging to that specific scene.

By visually inspecting the bar plot, viewers can gain insights into the dataset's composition and the relative abundance of different landing scene categories. This visualization helps researchers and practitioners understand the dataset's diversity and balance, which is crucial for training machine learning models effectively. Furthermore, it facilitates the identification of any class imbalances or biases present in the dataset, informing potential strategies for data augmentation or class weighting during model training.

Overall, the Seaborn and Matplotlib visualization provides a clear and intuitive representation of the distribution of landing scene images, aiding in the exploratory analysis and interpretation of the dataset's characteristics.

f) Feature Selection:

Feature selection in landing scene recognition involves identifying the most informative visual attributes essential for accurate classification while reducing computational complexity. This process entails leveraging domain knowledge to pinpoint relevant features like texture, color, and spatial relationships extracted from images using techniques such as Histogram of Oriented Gradients (HOG) or deep learning-based methods like convolutional neural networks (CNNs). Additionally, feature selection methods like Principal Component Analysis (PCA) or Recursive Feature Elimination (RFE) can automatically identify and retain the most discriminative features, enhancing model interpretability and generalization performance while mitigating the curse of dimensionality.

By selecting pertinent features and discarding redundant ones, feature selection optimizes model efficiency and robustness, crucial for the deployment of reliable autonomous drone landing systems. This approach not only streamlines computational resources but also improves the model's ability to accurately classify diverse landing environments, contributing to enhanced safety and efficiency in drone operations.

g) Training & Testing:

ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

In the training phase, the LandingScenes-7 dataset is split into training and testing sets to facilitate model development and evaluation. Two deep learning models, ResNext50 and ResNet50, are trained using the training data. These models leverage convolutional neural networks (CNNs), known for their effectiveness in extracting hierarchical features from images, and the ADAM optimizer, which adaptively adjusts learning rates for efficient parameter updates.

During training, the models iteratively process batches of training samples, adjusting their parameters to minimize the prediction error. The training process involves forward propagation to compute predictions, followed by backward propagation to calculate gradients and update model weights accordingly. This iterative optimization process continues until the models converge to a state where further training does not significantly improve performance on the training data.

Following training, the trained models are evaluated using the testing set to assess their performance in accurately recognizing landing scenes. Performance metrics such as accuracy, precision, recall, and F1 score are computed to quantify the models' effectiveness in classifying landing scene images. This evaluation phase helps determine the models' generalization capabilities and their suitability for real-world deployment in autonomous drone landing systems.

h) Algorithms:

ResNext50 with ADAM: ResNext50 extends the ResNet architecture by introducing the concept of cardinality, enabling the simultaneous learning of diverse features through multiple paths. This deep CNN is adept at capturing complex patterns and variations within images, making it well-suited for landing scene recognition. Paired with the ADAM[18] optimizer, ResNext50 dynamically adjusts its learning rates, facilitating efficient convergence during training. Its ability to grasp intricate details makes ResNext50 with ADAM a compelling choice for achieving high accuracy in classifying diverse landing environments.

ResNet50 with ADAM: ResNet50 leverages residual learning to effectively train very deep neural networks. Residual connections enable the learning of residual functions, addressing issues like the vanishing gradient problem. With its deep architecture, ResNet50 captures hierarchical features essential for image recognition tasks, including landing scene recognition. By utilizing the ADAM optimizer, ResNet50[12] adjusts learning rates based on the optimization landscape, contributing to efficient training. In the context of landing scene recognition, ResNet50 with ADAM emerges as a potent combination for robustly learning and classifying intricate scene patterns.

ResNext50+Random Forest: In the ResNext50+Random Forest approach, features extracted by the ResNext50 model are passed to a Random Forest classifier for final decision making. ResNext50[13], with its cardinality concept and deep architecture, extracts rich features from landing scene images. These features are then fed into the Random Forest classifier, which utilizes an ensemble of decision trees to make predictions. This combination harnesses the strengths of both deep learning and traditional machine learning methods, leveraging the powerful feature extraction capabilities of ResNext50

INTERNATIONAL JOURNAL OF
PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

and the robustness of Random Forest for classification tasks. In the domain of landing scene recognition, the ResNext50+Random Forest ensemble offers a versatile and effective solution for accurately classifying diverse landing environments.

4. EXPERIMENTAL RESULTS

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

Accuracy = TP + TN TP + TN + FP + FN.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

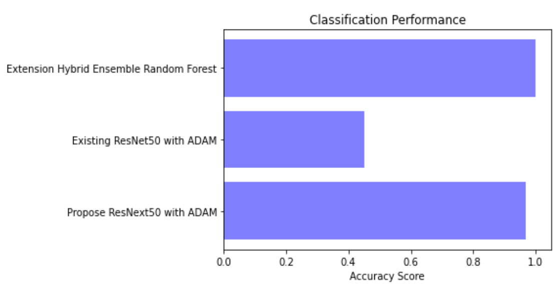

Fig 3 Accuracy Comparison Graph

F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

$$\textbf{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

$$\textbf{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

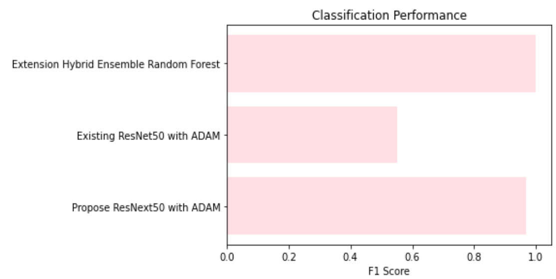

Fig 4 F1 Score Comparison Graph

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\text{Precision} = \frac{\textit{True Positive}}{\textit{True Positive} + \textit{False Positive}}$$

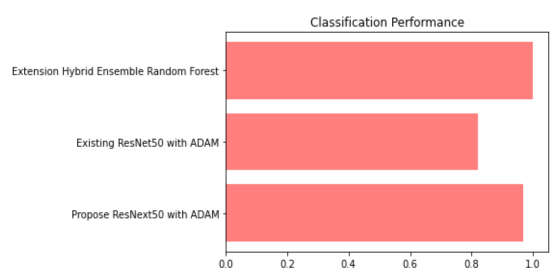

Fig 5 Precision Comparison Graph

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{TP}{TP + FN}$$

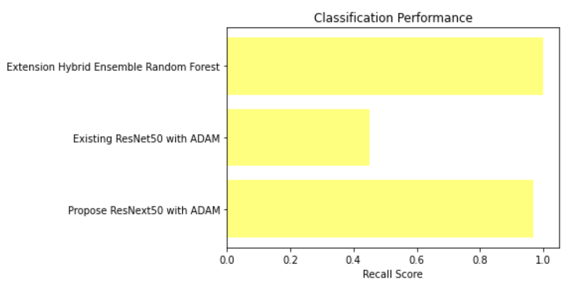

Fig 6 Recall Comparison Graph

	Algorithm Name	Precision	Recall	FScore	Accuracy
0	Existing ResNet50 Adam	97.107304	97.130906	97.09922	96.699029
1	Propose ResNext50 Adam	32.062088	46.533165	35.19191	45.048544
2	**Extension ResNext50 + Random Forest Hybrid Model**	**100.000000**	**100.000000**	**100.00000**	**100.000000**

Fig 7 Performance Evaluation Table

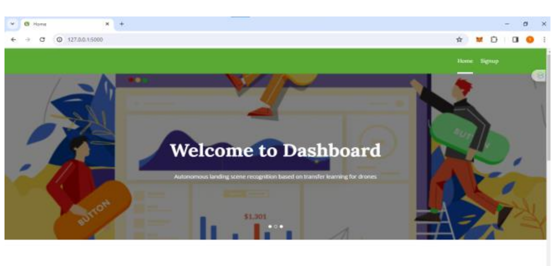

Fig 8 Home Page

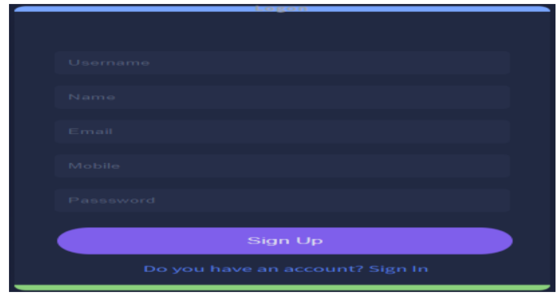

Fig 9 Registration Page

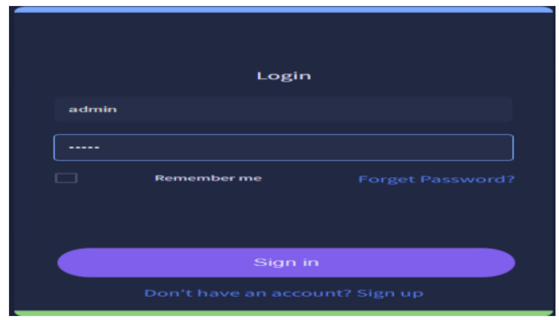

Fig 10 Login Page

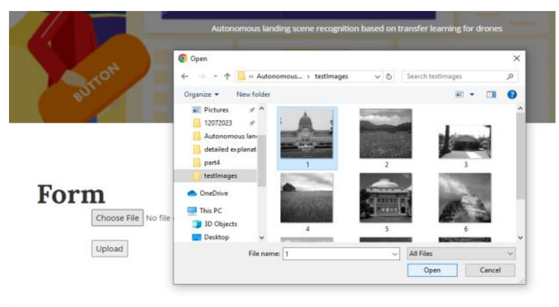

Fig 11 Upload Input Image

Fig 12 Final Outcome

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST

ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

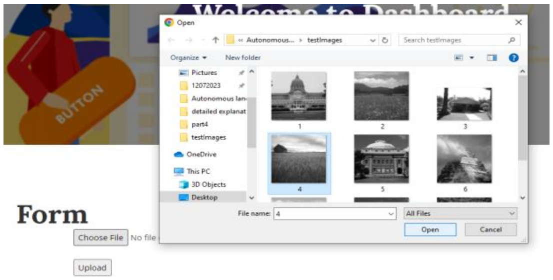

Fig 13 Upload Input Image

Fig 14 Predicted Results

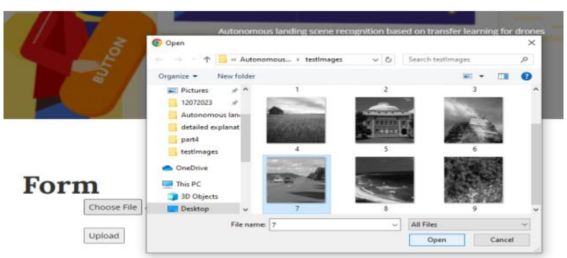

Fig 15 Upload Input Image

Fig 16 Final Outcome

5. CONCLUSION

In conclusion, this paper presents a significant advancement in drone technology through the development of an autonomous landing scene recognition system. By leveraging state-of-the-art artificial intelligence techniques, including ResNext50[13] and ResNet50[12] models, the project enhances the autonomy and safety of drone operations. The establishment of the LandingScenes-7 dataset addresses the need for standardized datasets in landing scene recognition research. Moreover, the integration of thresholding methods for novelty detection paves the way for real-time landing scene categorization, ensuring drones can adapt to unforeseen environments.

The project's outcomes have far-reaching implications for various industries utilizing drone services. By enabling safer and more accurate drone operations, particularly in emergency situations, the developed models contribute to enhanced efficiency and reliability in applications such as surveillance, news reporting, and logistics. The introduction of a hybrid model, combining ResNext50 and Ensemble Random Forest, showcases a novel approach to improving prediction accuracy, further bolstering the effectiveness of the landing scene recognition system.

6. FUTURE SCOPE

Looking ahead, there are several avenues for future research and development in the field of autonomous landing scene recognition for drones. One potential direction involves exploring additional deep learning architectures and ensemble techniques to further improve model performance and robustness. Moreover, expanding the dataset to include a more

ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

diverse range of landing scenes and environmental conditions could enhance the models' ability to generalize to real-world scenarios.

Furthermore, incorporating real-time environmental sensing capabilities, such as LiDAR or radar sensors, could provide valuable complementary information to improve landing scene recognition accuracy, particularly in challenging weather conditions or low visibility environments. Additionally, integrating advanced decision-making algorithms and reinforcement learning techniques could enable drones to adapt their landing strategies dynamically based on environmental cues and mission objectives.

Overall, the future of autonomous landing scene recognition holds promise for advancing drone technology and enhancing safety and efficiency in a wide range of applications. Continued research and innovation in this area will undoubtedly contribute to the continued evolution of drone capabilities and their integration into various industries and societal contexts.

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 images. Proc. 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 vehicle. Proc. 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 K. Improving object recognition with the l-channel. Pattern Recognition, 2016, 49: 187–197.

[7] ZHOU B L, LAPEDRIZA A, KHOSLA A, et al. Places: a 10 million image database for scene recognition. IEEE Trans. On Pattern Analysis and Machine Intelligence, 2018, 40(6): 1452–1464.

[8]XIAO J X, HAYS J, EHINGER K A, et al. SUN database: large-scale scene recognition from abbey to zoo. Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2010: 3485–3492.

[9]XIAO J X, EHINGER K A, HAYS J, et al. SUN database: exploring a large collection of scene categories. International Journal of Computer Vision, 2016, 119(1): 3–22.

INTERNATIONAL JOURNAL OF PURE AND APPLIED SCIENCE & TECHNOLOGY
IJPAST
ISSN 2229-6107 www.ijpast.in
Vol 14,Issuse 2.May 2024

[10]PATTERSON G, XU C, SU H, et al. The SUN attributedatabase: beyond categories for deeper scene understanding.International 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 al. Deep residual learning for image recognition. Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770–778.

[13]XIE S N, GIRSHICK R, DOLLAR P, et al. Aggregated residual transformations for deep neural networks. Proc. Of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2017: 5987–5995.

[14]XIE L, LEE F F, LIU L, et al. Scene recognition: a comprehensive survey. Pattern Recognition, 2020, 102: 107205.

[15]NEYSHABUR B, SEDGHI H, ZHANG C Y. What is beingtransferred in transfer learning? Proc. of the 34th Conferenceon Neural Information Processing Systems, 2020: 512–523.

[16]WU X Z, LIU R Z, YANG H Q, et al. An Xception basedconvolutional neural network for scene image classificationwith transfer learning. Proc. of the 2nd International Conference on Information Technology and Computer Application,2020: 262–267.

[17]STEGER C, UIRICH M, WIEDEMANN C. Machine visionalgorithms and applications. Berlin: Wiley-VCH, 2018.

[18]KINGMA D P, BA J L. Adam: a method for stochastic optimization. Proc. of the 3rd International Conference on Learning Representations. https://arxiv.org/abs/1412.6980.

[19]QIAN N. On the momentum term in gradient descent learning algorithms. Neural Networks, 1999, 12(1): 145–151.

[20] DUCHI J, HAZAN E, SINGER Y. Adaptive subgradient
methods for online learning and stochastic optimi zation.
Journal of Machine Learning Research, 2011, 12(61): 2121–2159.